



SOVAT: An OLAP/GIS system for Environmental Health Tracking

Bambang Parmanto, PhD

The Challenge

- Environmental Health Tracking requires numerical-spatial analysis: map, charts, computation
- Environmental Health Tracking involves large, multidimensional data sets
- No DSS that is effectively able to support this
- Ease of use and cost-effective tool is needed for widespread tracking to be adopted



Inspiration & Historical Perspectives: John Snow

- Cholera Outbreak in London in 1854
- Used numerical-spatial problem solving to discover source of the cholera outbreak
- Could we do better than John Snow of a century and a half ago?



John Snow
pinpointed the
source of
cholera
outbreak in
London, 1854



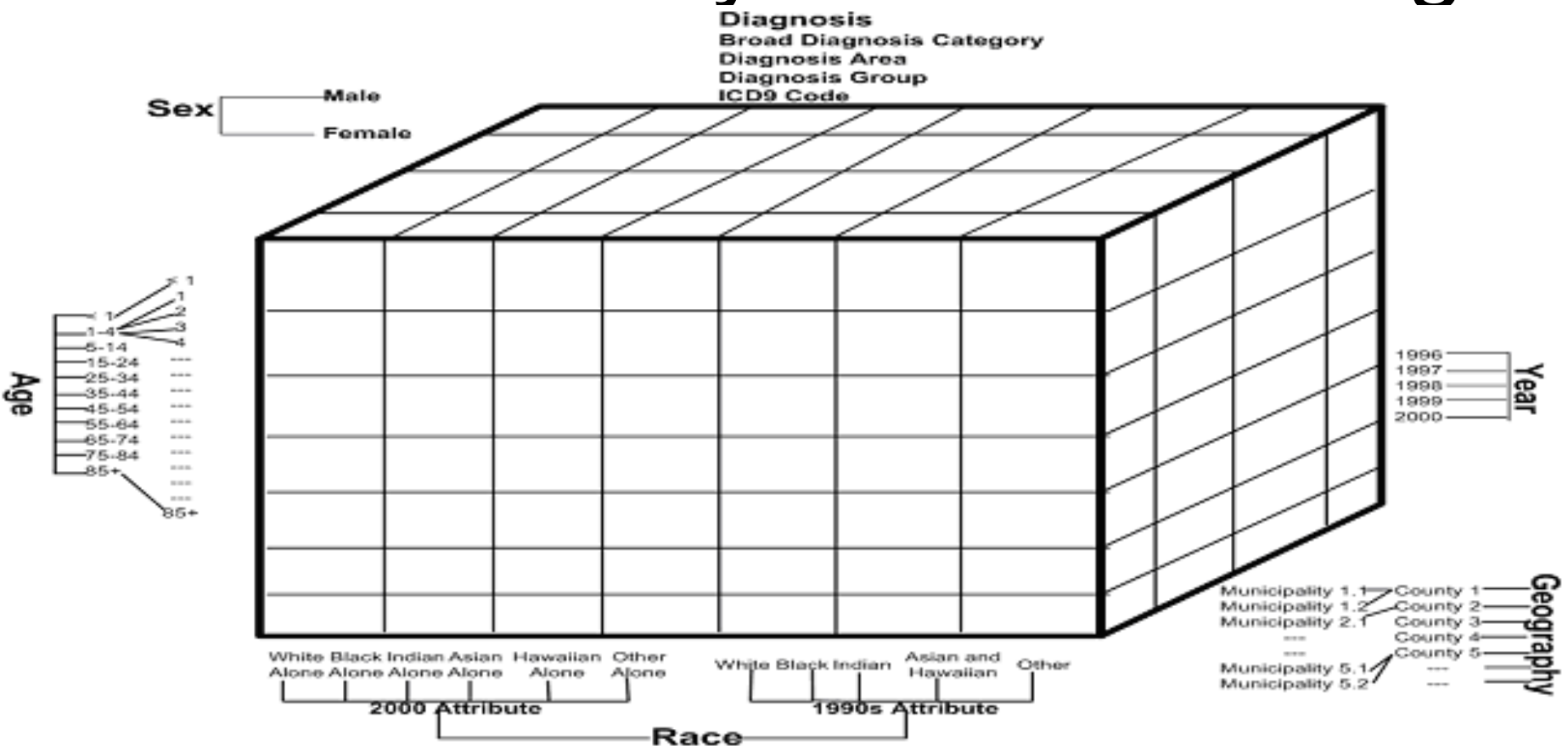
Numerical-Spatial Routine: Example of Disease mapping

	Component	Type
1	Identify geographic area	Spatial
2	Exposure or hazard analysis/calculation of the geographic area	Numerical
3	Identify comparison region	Spatial
4	Analysis of cancer/heart disease incidence	Numerical
5	View in charts/reports	Numerical
6	View Maps	Spatial

OLAP & GIS

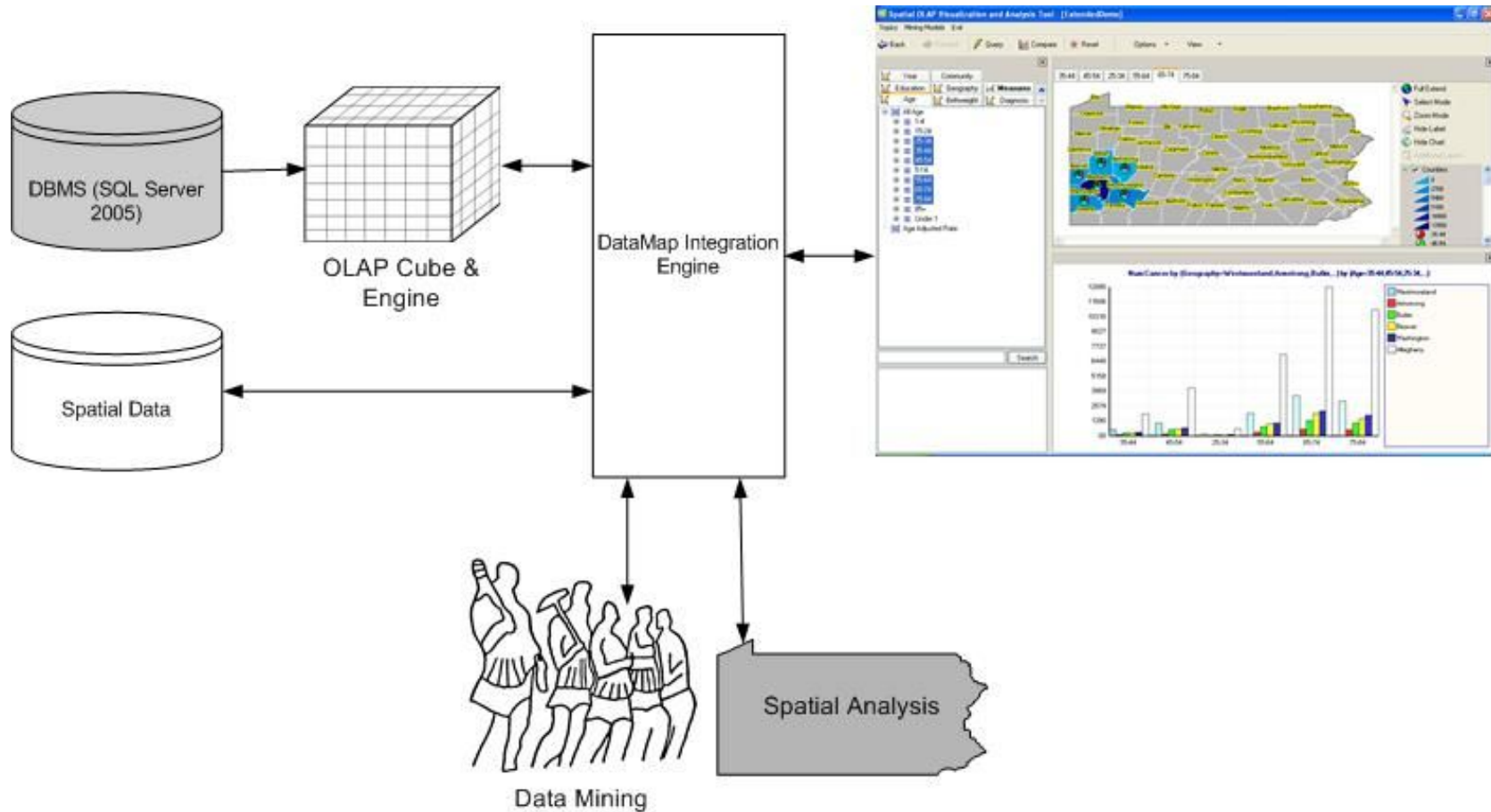
- OLAP is very attractive for environmental health analyses: handle large amount of data, multidimensional, fast retrieval
- OLAP enhances numerical components
- GIS is central to environmental health tracking: it needs to have the capability to do spatial analysis:
 - Network Analysis, Buffering, Clipping, Geo-coding
- Integrated system: OLAP and GIS needs to be integrated into an easy-to-use system.

Online Analytical Processing



Functions: Drill-up/Drill-down, Slice/Dice, Pivot

SOVAT Architecture

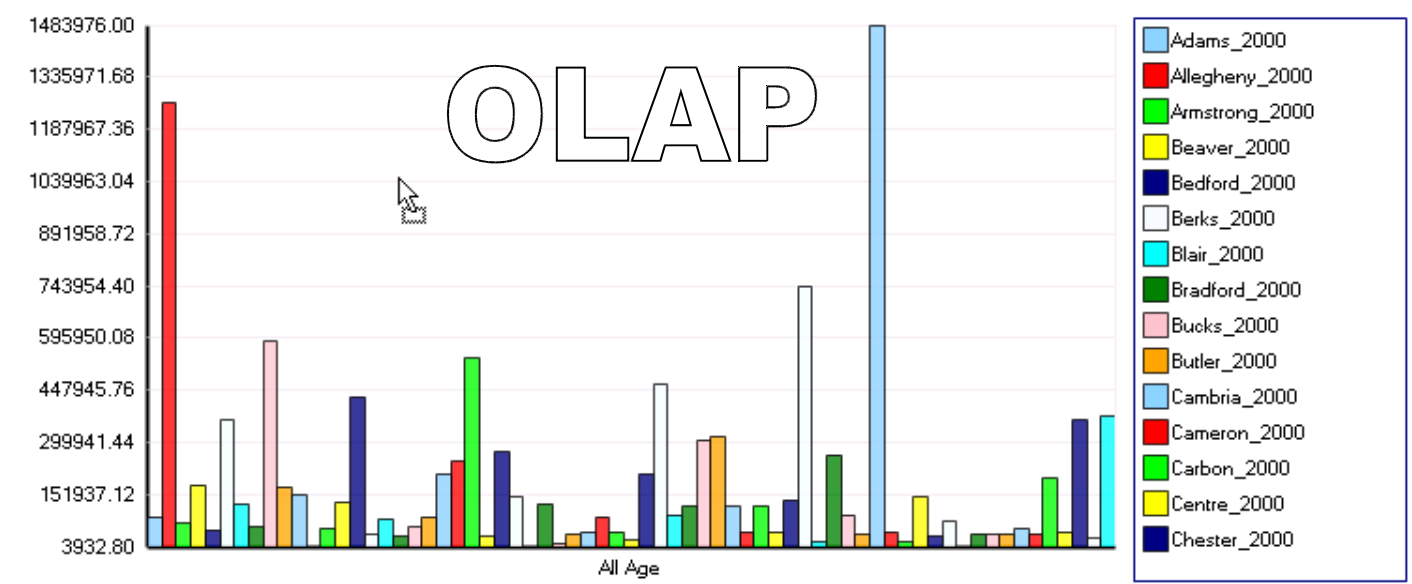


Dimensions

Age Birthweight Diagnosis

- All Diagnosis
 - Certain conditions originating
 - Complications of pregnancy
 - Complications mainly related to
 - Complications occurring during
 - Complications of the pregnancy
 - Ectopic and molar pregnancies
 - Normal delivery, and other
 - Congenital anomalies
 - Endocrine, nutritional and
 - External causes of injury and
 - Infectious and parasitic diseases
 - Injury and Poisoning
 - Mental disorders
 - Neoplasms
 - Symptoms, signs, and ill-defined
 - the circulatory system

Population



OLAP

Age Geography Year

Background

- Birthweight
- Diagnosis
- Education
- Measures
- Race
- Region
- Sex

Query

Top 5

Time Series

Reset

All Age_2000

☒ Counties

0 320000 630000 940000 1250000 1560000

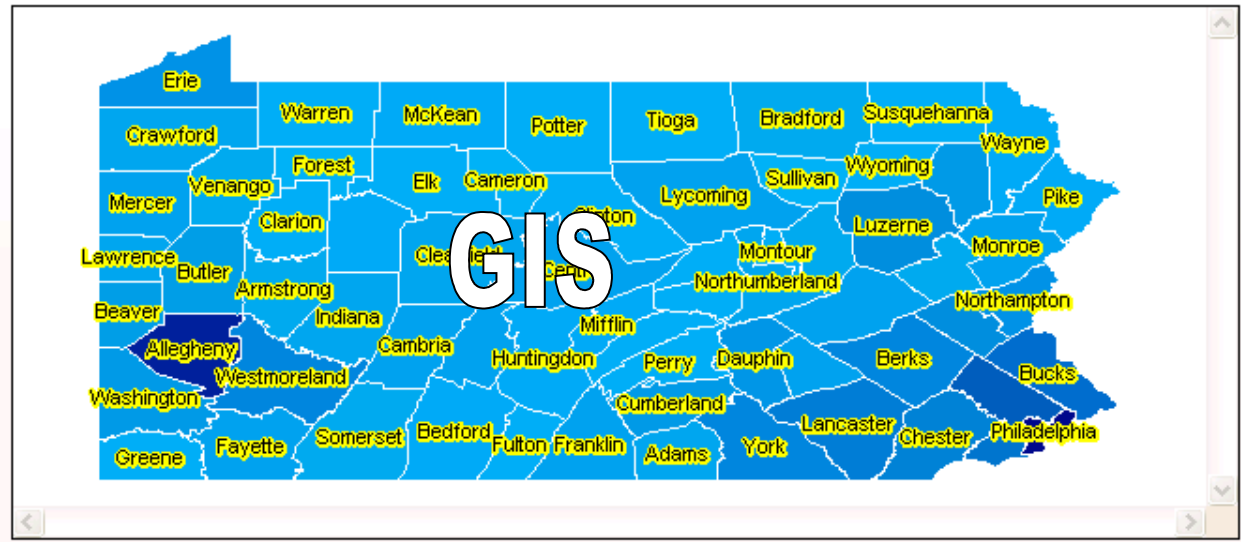
☒ buffer

Full Extend

Select Mode

Zoom Mode

Drag Mode



Dimensions

- Geography Measures Race
- um Births
 - um Inpatient
 - IS
 - atient Charges
 - um Outpatient
 - utpatient Charges
 - um Cancer
 - um Deaths
 - opulation
 - IS/1,000
 - atient/1,000**
 - utpatient/1,000
 - ancer Incidence/100,000
 - death/100,000
 - death/100,000 (Smoothed)
 - th/100,000 (Smoothed)
 - ancer Incident/100,000 (Smoothed)

Row Column

Age Geography

Year

Background

- Birthweight
- Diagnosis
- Education
- Measures
- Race
- Region
- Sex

Query

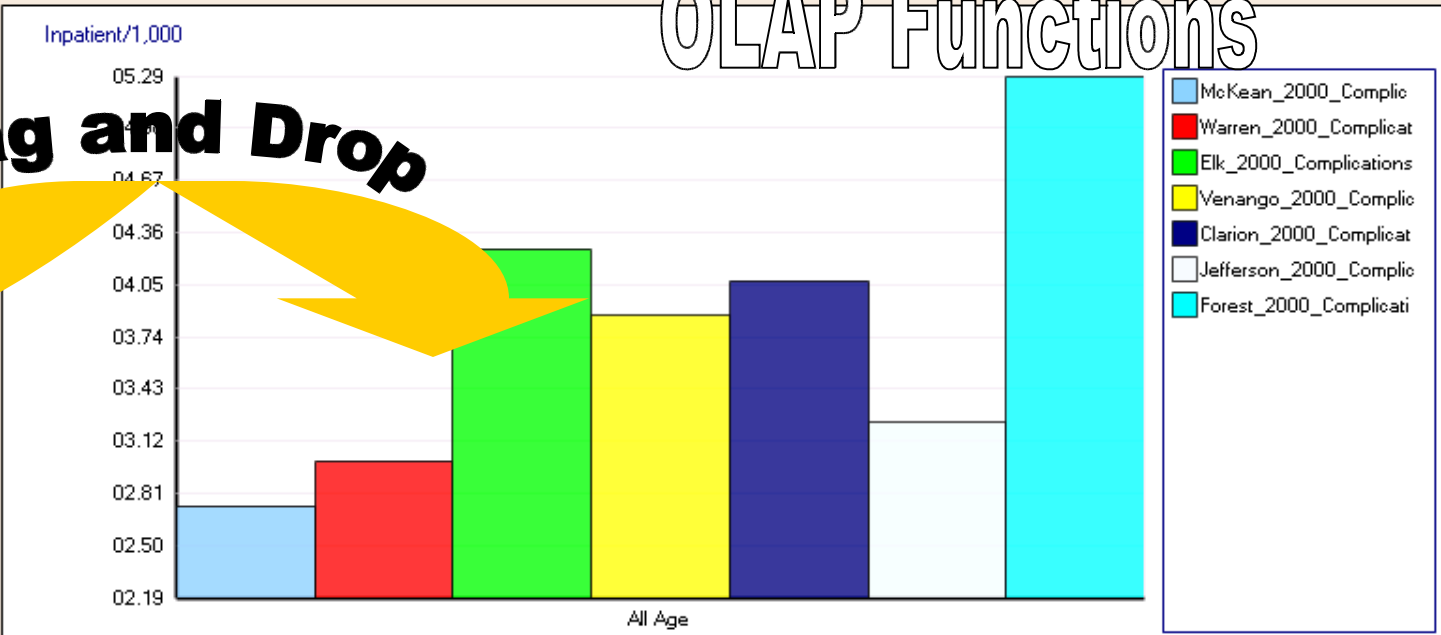
Top 5

Time Series

Reset

OLAP Functions

Drag and Drop



OLAP-GIS Functions

All Age_2000_Complications occurring mainly in the course of labor and delivery

☐ ☒ Counties

- 2.6
- 3.2
- 3.8
- 4.4
- 5.0
- 5.6

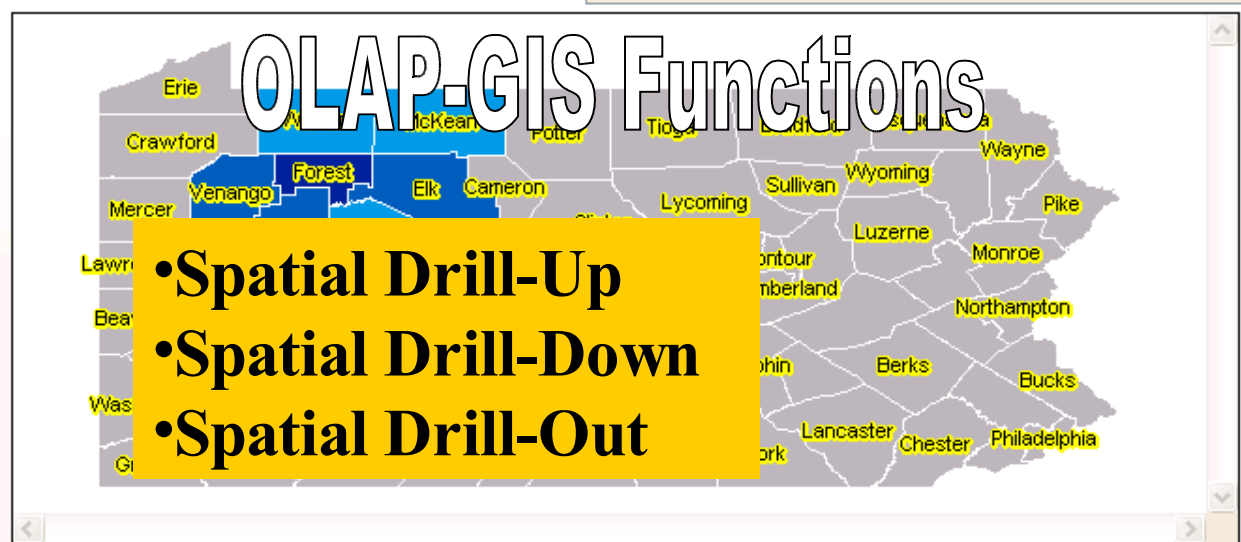
☐ ☒ buffer

Full Extend

Select Mode

Zoom Mode

Drag Mode

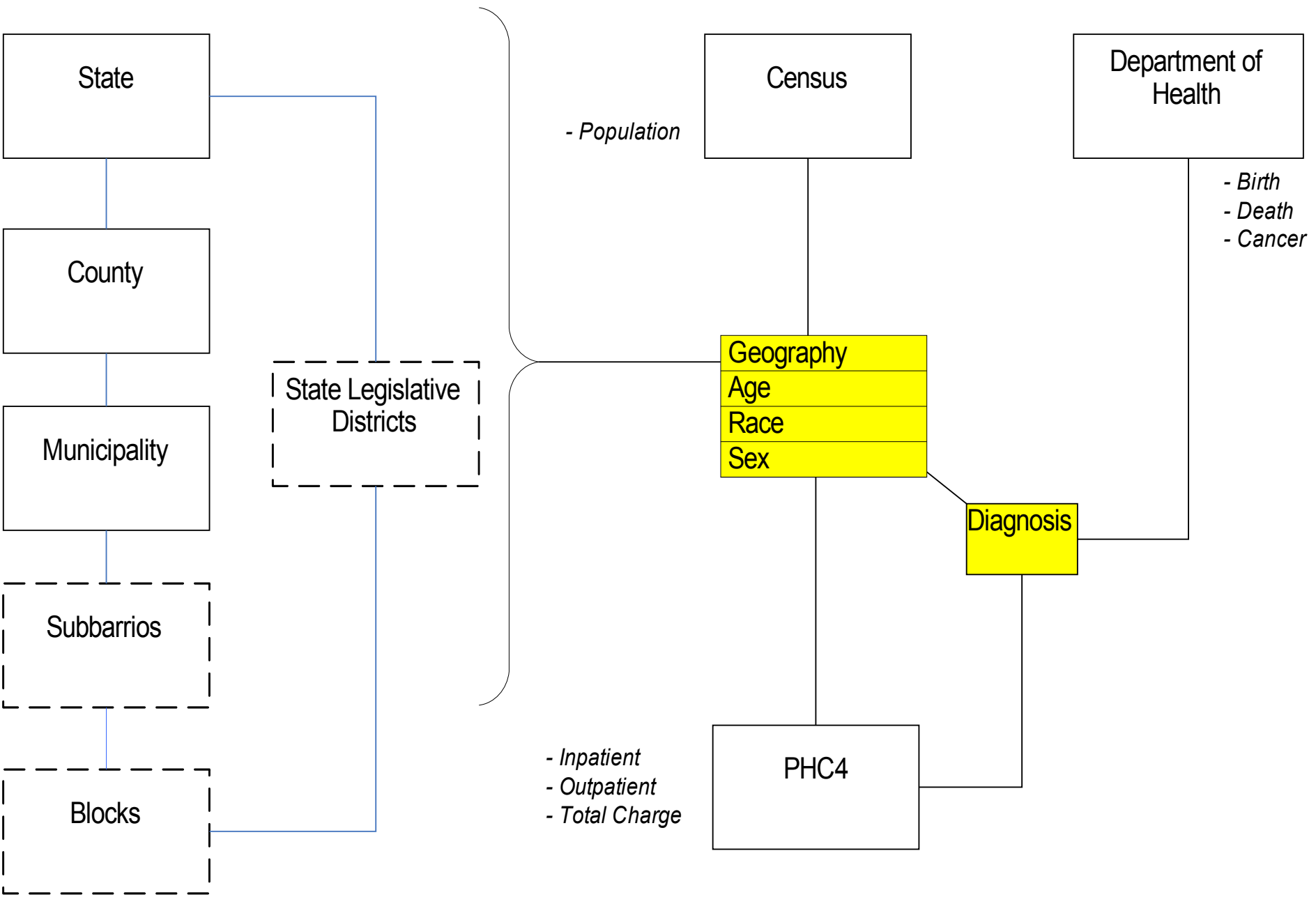


- Spatial Drill-Up
- Spatial Drill-Down
- Spatial Drill-Out



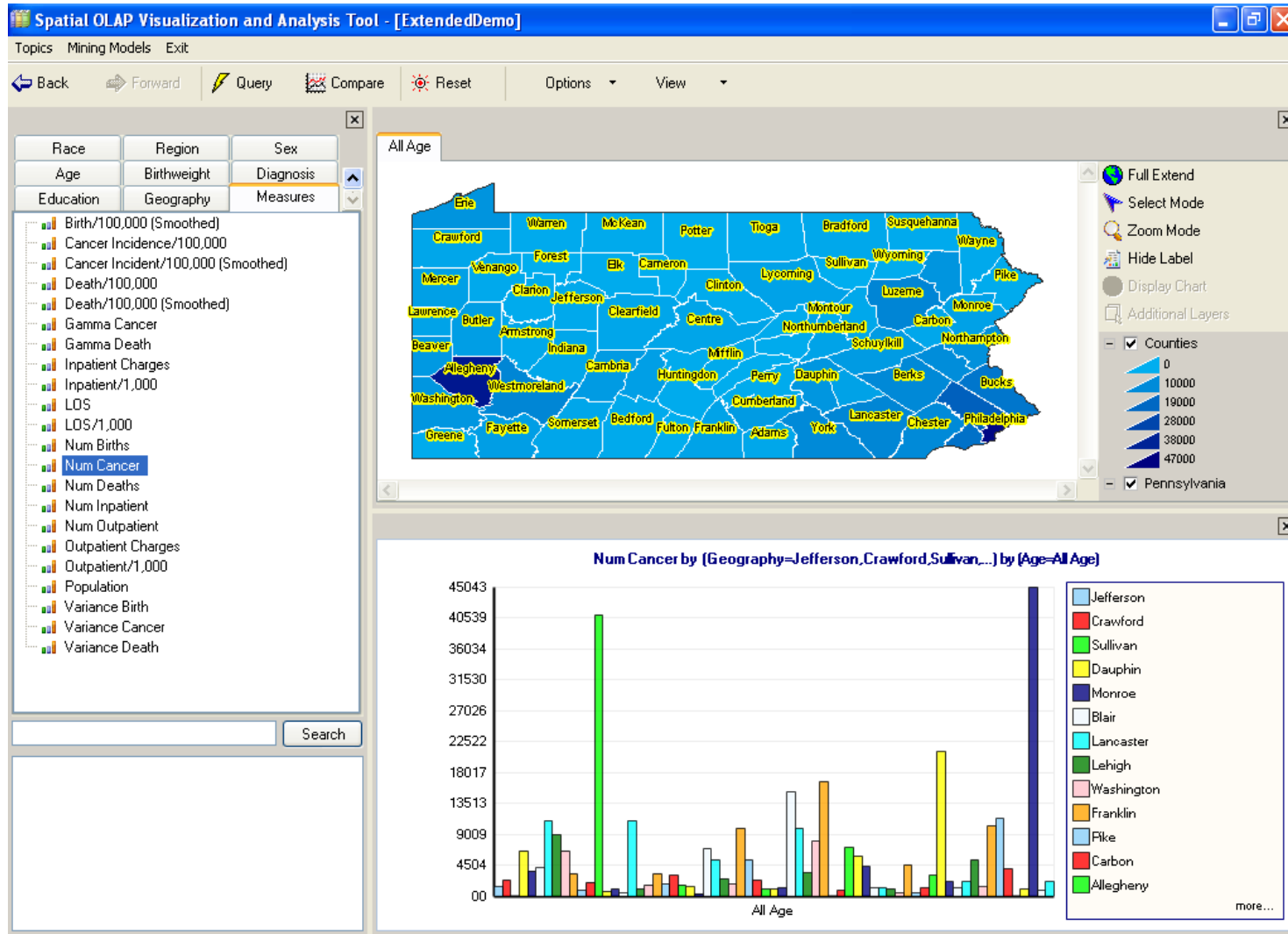
Example of Pennsylvania Data

- Spatial Data of Pennsylvania
 - US Census Bureau .shp files
- Numerical data from 3 sources:
 1. U.S. Census Bureau
 - Population & Socioeconomic
 3. PA Department of Health
 - Cancer incidence, death, birth
 5. PA Healthcare Cost Containment Council
 - Inpatient & Outpatient Healthcare Utilization





Features: Default presentation, Time Series, Top 5 for Certain Disease in PA



Dimensions

Region Sex Year Commu < >

- All Year
 - 1996
 - 1997
 - 1998
 - 1999
 - 2000

Row

Column

Geography

Age

Background

- Birthweight
- Diagnosis
- Education
- Measures
- Race
- Region
- Sex
- Year

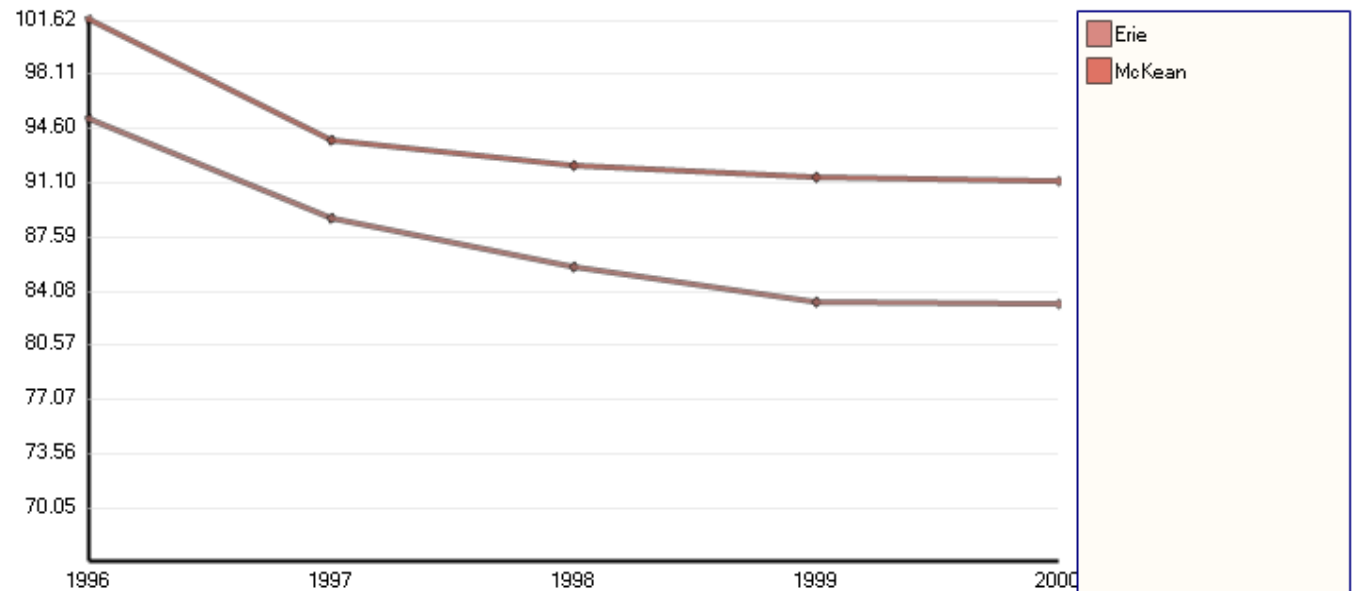
Query

Top 5

Compare

Reset

Inpatient/1,000 (Geography=Erie,McKean) (Year=1996,1997,1998,1999,2000)



All Age_Inpatient/1,000_1996

All Age_Inpatient/1,000_1997

All Age_Inpatient/1,000_1998

All Age_Inpatient/1,000_1999

All Age_Inpatient/1,000_2000 < >

☒ Counties

- 79
- 85
- 90
- 96
- 101
- 107

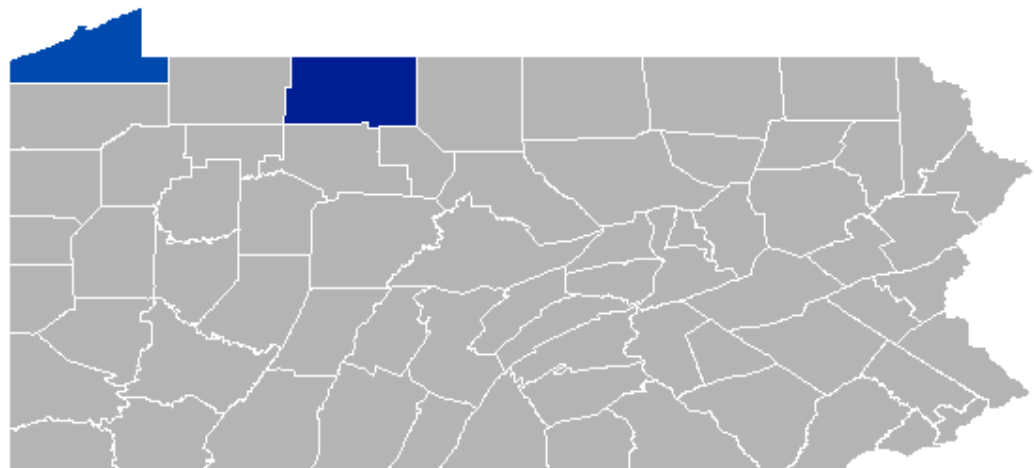
☒ buffer

Full Extend

Select Mode

Zoom Mode

Drag Mode



Dimensions

Age Birthweight **Diagnosis**

- Complications of pregnancy, childbirth or neonatal anomalies
- Endocrine, nutritional and metabolic
- External causes of injury and poisoning
- Infectious and parasitic diseases
- Injury and Poisoning
- Mental disorders
- Neoplasms**
 - Benign neoplasms
 - Carcinoma in situ
 - Malignant neoplasm of bone, connective tissue or cartilage
 - Malignant melanoma of skin
 - Malignant neoplasm of bone
 - Malignant neoplasm of connective tissue
 - Malignant neoplasm of female
 - Malignant neoplasm of male
 - Other malignant neoplasm of

Row Column

Age Geography

Background

- Birthweight
- Diagnosis
- Education
- Measures
- Race
- Region
- Sex
- Year

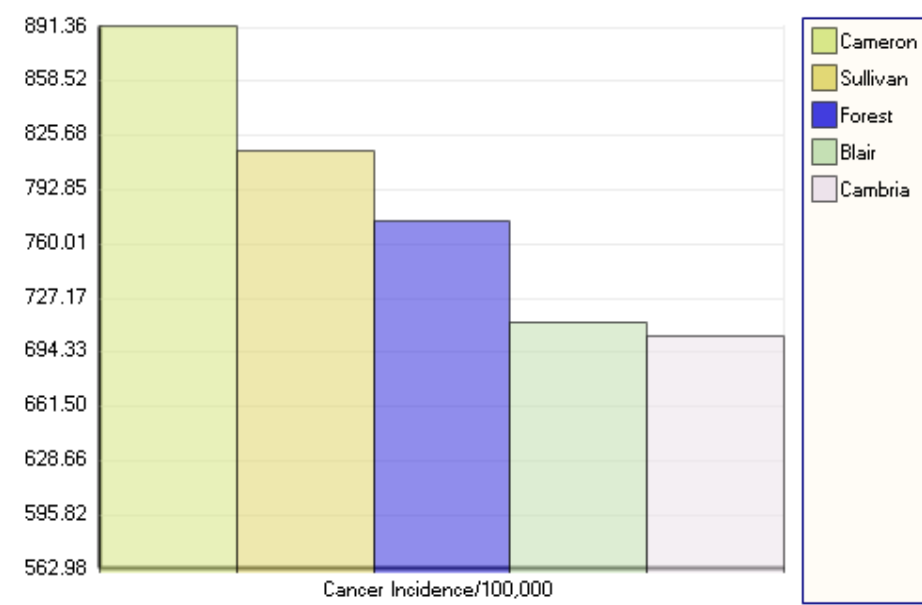
Query

Top 5

Compare

Reset

Cancer Incidence/100,000 - TOP 5 Geography (Year=2000) (Geography=Pennsylvania) (Diagnosis=Neoplasms)



Cancer Incidence/100,000

☐ ☒ Counties

- 670
- 720
- 780
- 830
- 880
- 940

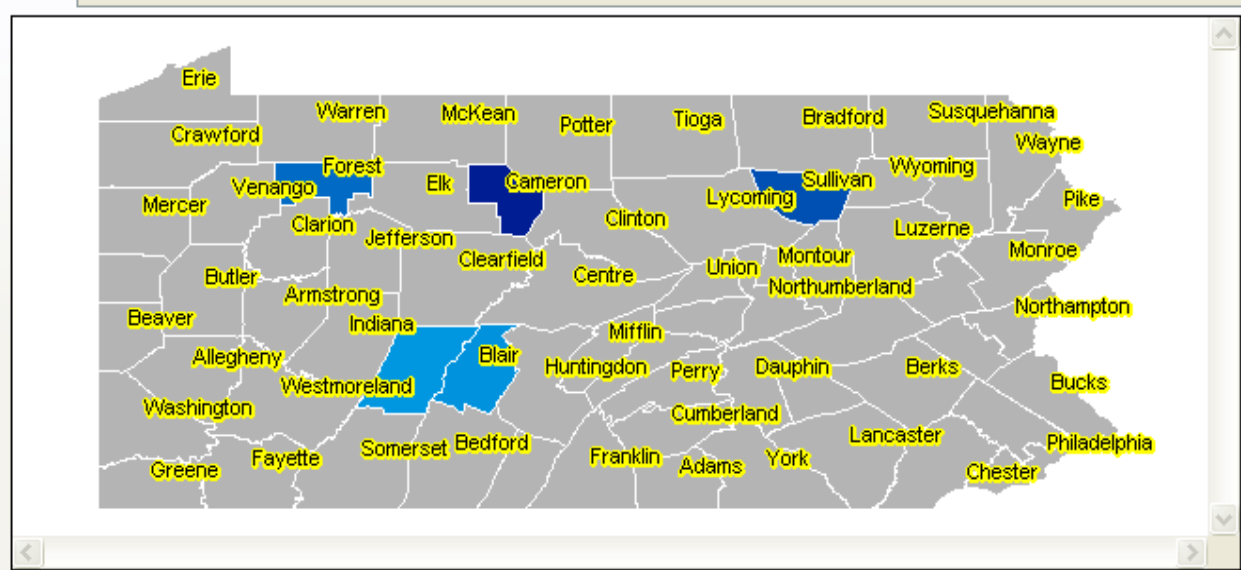
☐ ☒ buffer

Full Extend

Select Mode

Zoom Mode

Drag Mode

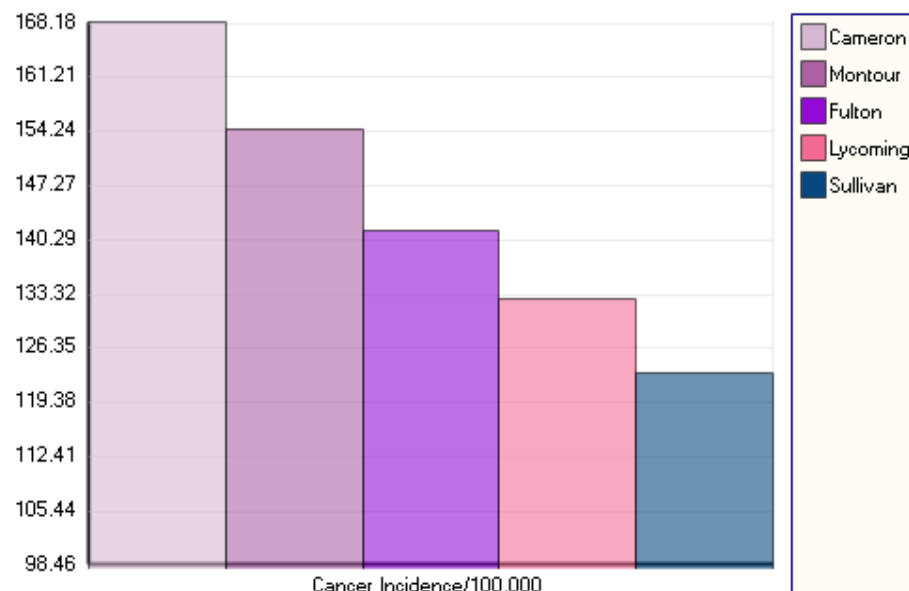


Dimensions

Age Birthweight **Diagnosis**

omplications of pregnancy, childbir
onegenital anomalies
ndocrine, nutritional and metabolic
xternal causes of injury and poison
fectious and parasitic diseases
jury and Poisoning
ental disorders
eoplasms
Benign neoplasms
Carcinoma in situ
Malignant neoplasm of bone, co
Malignant melanoma of skin
Malignant neoplasm of bone
Malignant neoplasm of conn
Malignant neoplasm of fema
Malignant neoplasm of male
Other malignant neoplasm o

Cancer Incidence/100,000 - TOP 5 Geography (Year=2000) (Geography=Pennsylvania) (Diagnosis=Malignant neoplasm of female breast)



Row

Column

Age

Geography

Background

Birthweight
Diagnosis
Education
Measures
Race
Region
Sex
Year

Query

Top 5

Compare

Reset

Cancer Incidence/100,000

☒ Counties



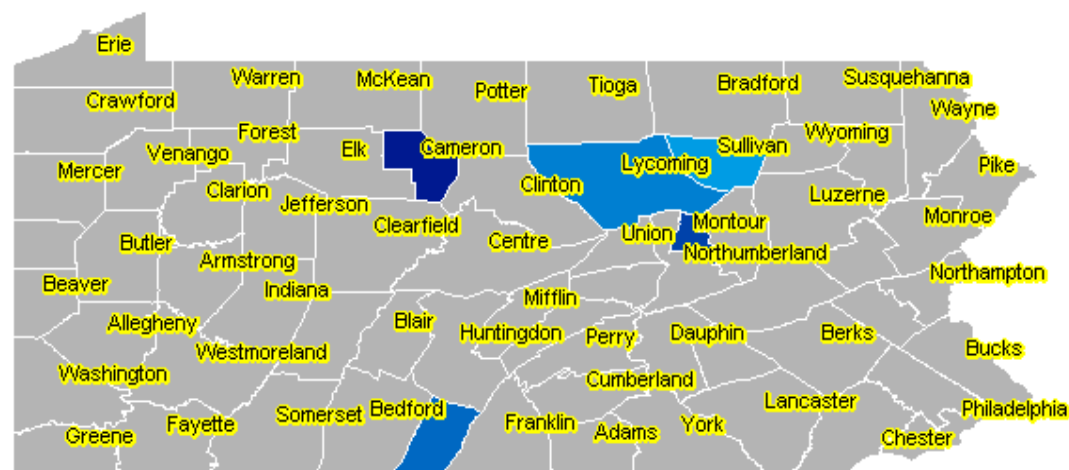
☒ buffer

Full Extend

Select Mode

Zoom Mode

Drag Mode



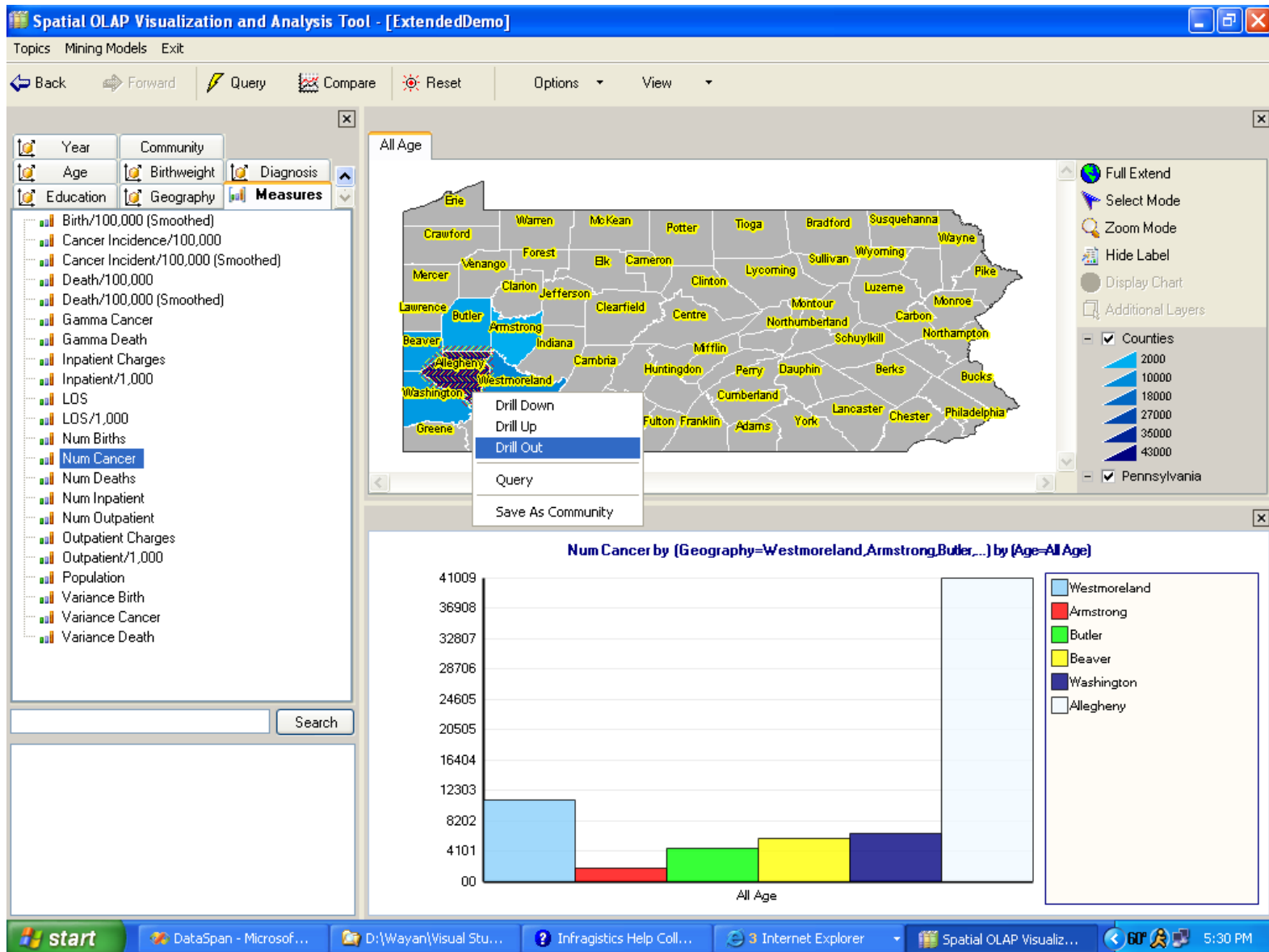


Problem Solving Example #2

“Compare Allegheny County’s cancer incidence rate against it’s bordering counties.”



Features: Comparison and Border Analysis





Spatial OLAP Visualization and Analysis Tool - [ExtendedDemo]

Topics Mining Models Exit

Back Forward Query Compare Reset

Options View

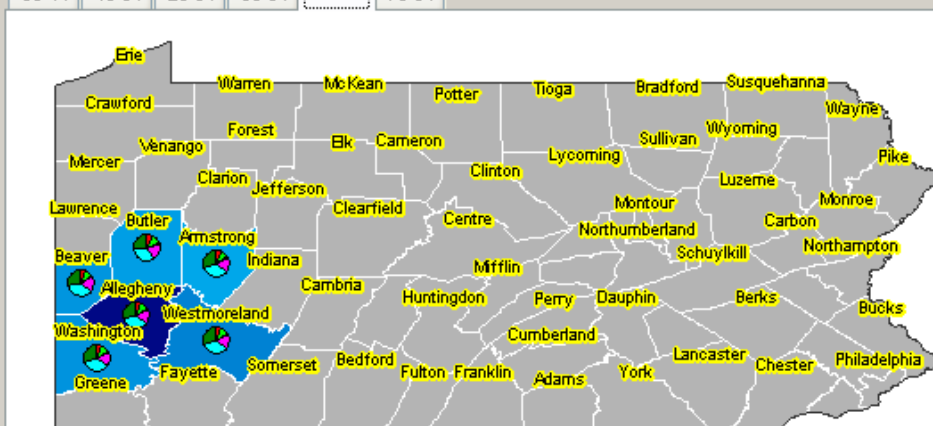
Race	Region	Sex
Education	Geography	Measures
Age	Birthweight	Diagnosis

All Age

- 1-4
- 15-24
- 25-34
- 35-44
- 45-54
- 5-14
- 55-64
- 65-74
- 75-84
- 85+
- Under 1

Age Adjusted Rate

35-44 45-54 25-34 55-64 65-74 75-84



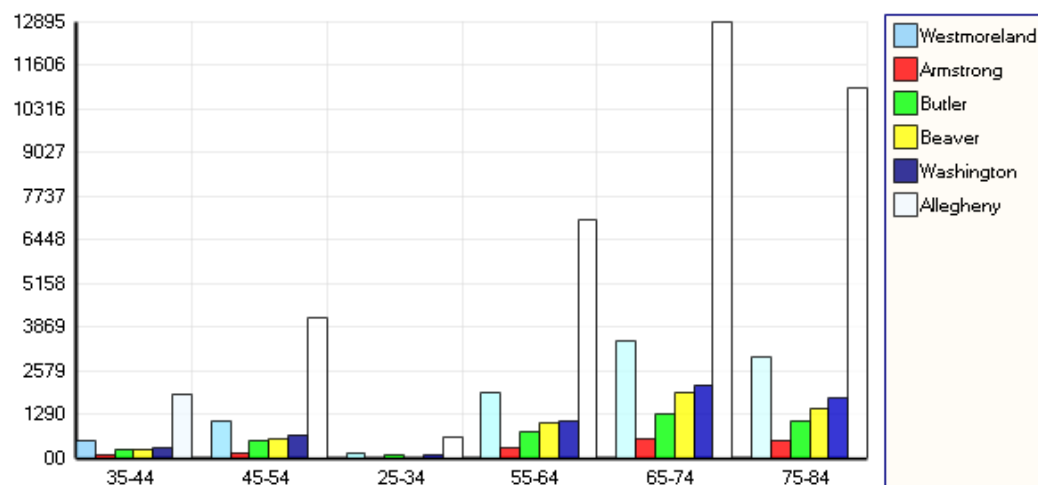
- Full Extend
- Select Mode
- Zoom Mode
- Hide Label
- Hide Chart

Additional Layers



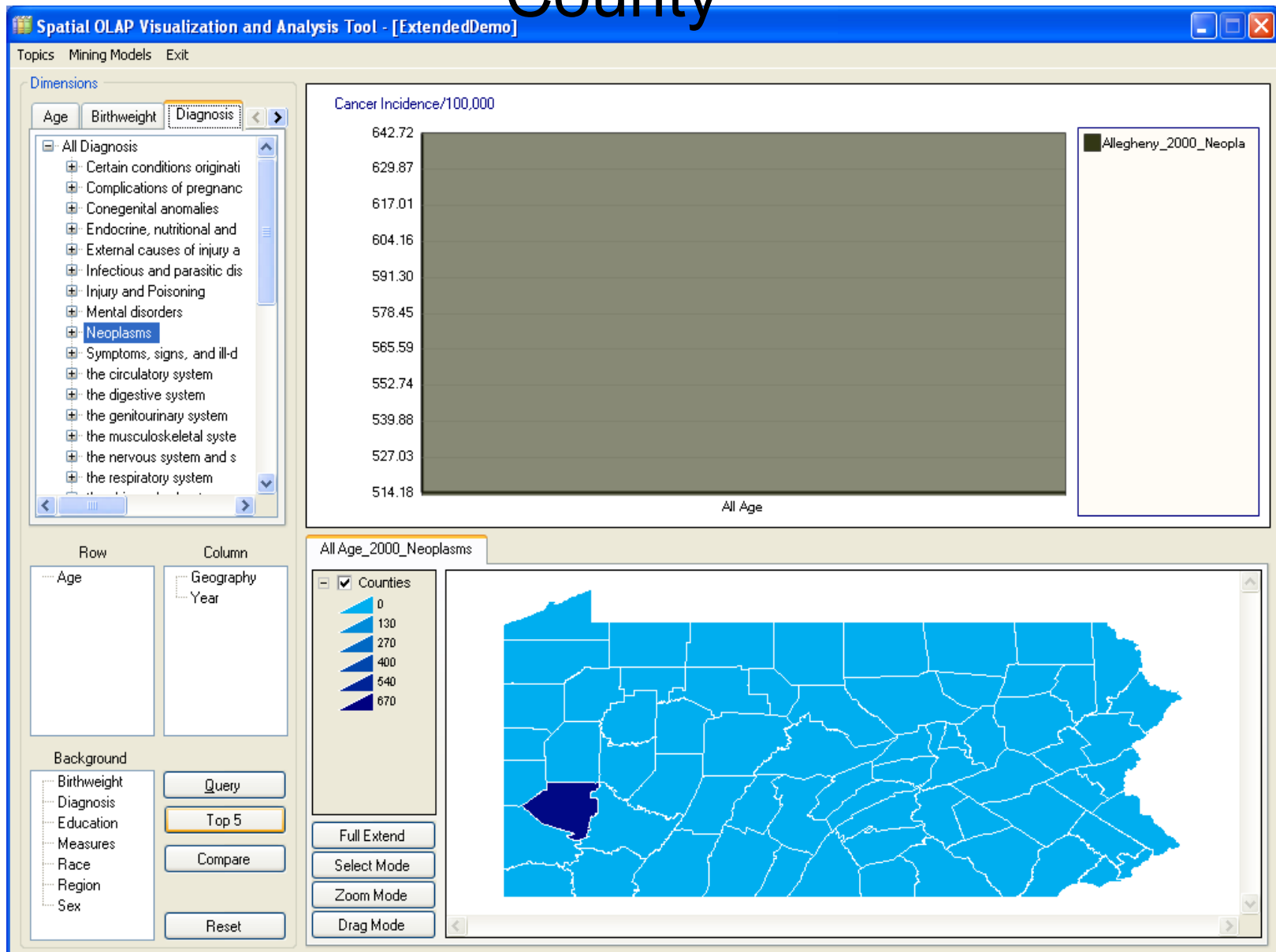
☒ Pennsylvania

Num Cancer by (Geography=Westmoreland,Armstrong,Butler,...) by (Age=35-44,45-54,25-34,...)





Features: Top 5 Cancer in Allegheny County



Dimensions

Diagnosis Education Geograp < >

All Diagnosis

- + Certain conditions originati
- + Complications of pregnanc
- + Congenital anomalies
- + Endocrine, nutritional and
- + External causes of injury a
- + Infectious and parasitic dis
- + Injury and Poisoning
- + Mental disorders
- + Neoplasms
- + Symptoms, signs, and ill-d
- + the circulatory system
- + the digestive system
- + the genitourinary system
- + the musculoskeletal syste
- + the nervous system and s
- + the respiratory system

Row

Geography

Column

Diagnosis

Background

- Birthweight
- Education
- Measures
- Race
- Region
- Sex
- Age
- Year

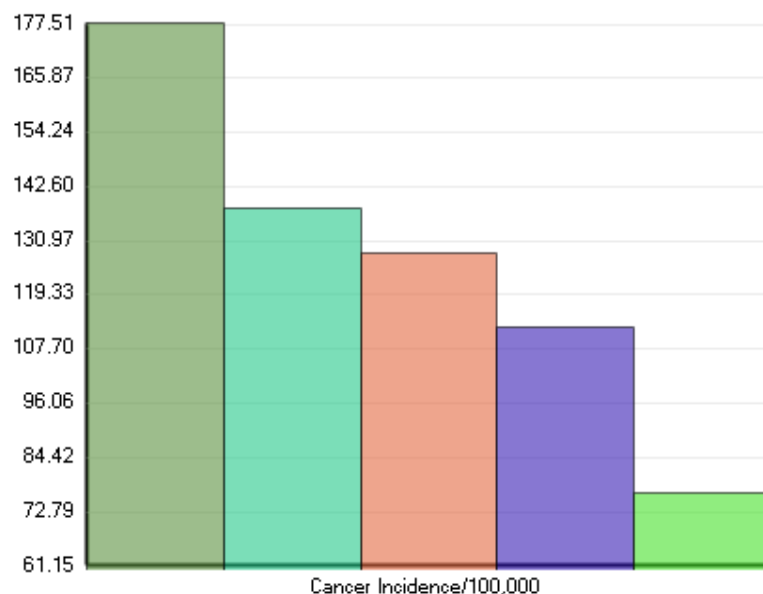
Query

Top 5

Compare

Reset

Cancer Incidence/100,000 - TOP 5 Diagnosis (Geography=Allegheny) (Diagnosis=Neoplasms)



- Malignant neoplasm of genitourinary organs
- Malignant neoplasm of digestive organs and peritoneum
- Malignant neoplasm of bone, connective tissue, skin, an
- Malignant neoplasm of respiratory and intrathoracic orga
- Malignant neoplasm of other and unspecified sites

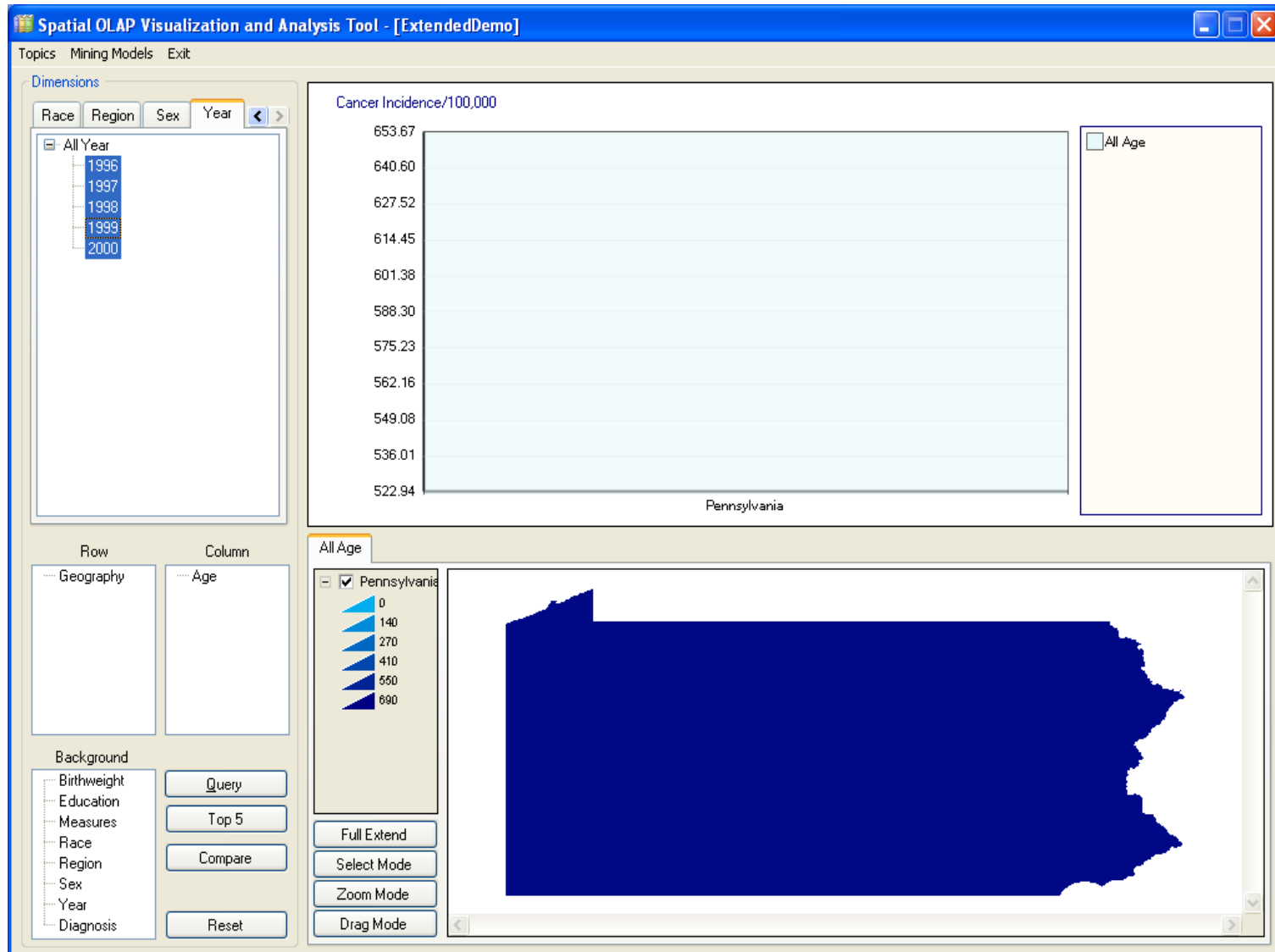


Problem Solving Example #4

“Compare Cancer Incidence of Allegheny County to Erie County from 1996-2000”



Features: Compare Cancer Incidence of Allegheny County to Erie County from 1996-2000



Dimensions

Education Geography Measure

- Clearfield
- Clinton
- Columbia
- Crawford
- Cumberland
- Dauphin
- Delaware
- Elk
- Erie
- Fayette
- Forest
- Franklin
- Fulton
- Greene
- Huntingdon
- Indiana
- Jefferson
- Juniata

Row

Column

Geography

Age

Background

- Birthweight
- Education
- Measures
- Race
- Region
- Sex
- Year
- Diagnosis

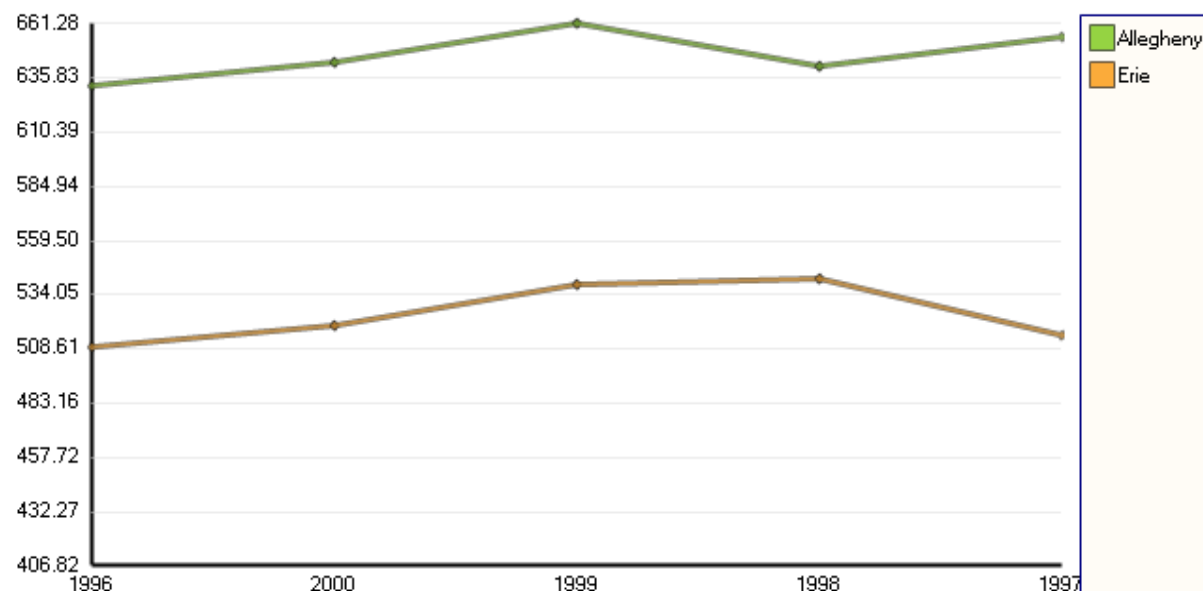
Query

Top 5

Compare

Reset

Cancer Incidence/100,000 (Geography=Allegheny,Erie) (Year=1996,2000,1999,1998,1997)



All Age_Cancer Incidence/100,000_1996

All Age_Cancer Incidence/100,000_2000

All Age_Cancer Incidence/100,000_1999

All Age_Cancer Incidence/100,000_1998

Counties

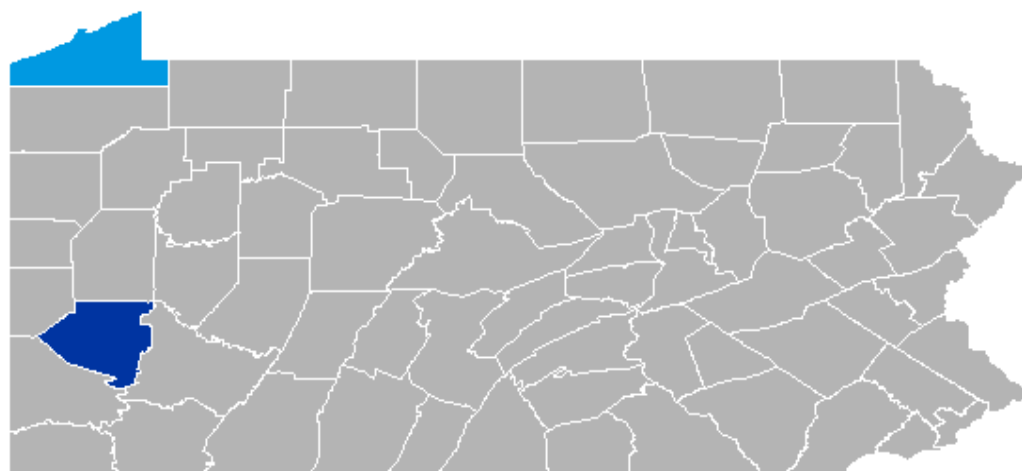
- 480
- 530
- 570
- 610
- 650
- 690

Full Extend

Select Mode

Zoom Mode

Drag Mode





Problem Solving Example

“Compare the 2000 death rates from ischemic heart disease of Northwestern PA vs. Northeastern PA”

Dimensions

Birthweight Diagnosis Educate < >

- + Complications of pregnanc
- + Congenital anomalies
- + Endocrine, nutritional and
- + External causes of injury a
- + Infectious and parasitic dis
- + Injury and Poisoning
- + Mental disorders
- + Neoplasms
- + Symptoms, signs, and ill-d
- + the circulatory system
 - + Acute rheumatic fever
 - + arteries, arterioles, an
 - + Cerebrovascular disea
 - + Chronic rheumatic hea
 - + Hypertensive disease
 - + Ischemic heart diseas
 - + Other forms of heart di

Row

Age

Column

Geography
Year

Background

Birthweight
Diagnosis
Education
Measures
Race
Region
Sex

Query

Top 5

Compare

Reset

PA Population

12138945.00
11896166.10
11653387.20
11410608.30
11167829.40
10925050.50
10682271.60
10439492.70
10196713.80
9953934.90
9711156.00

All Age

Pennsylvania_2000

All Age_2000

- ☒ Pennsylvania
 - 0
 - 2500000
 - 5100000
 - 7600000
 - 10200000
 - 12700000
- ☒ buffer

Full Extend

Select Mode

Zoom Mode

Drag Mode

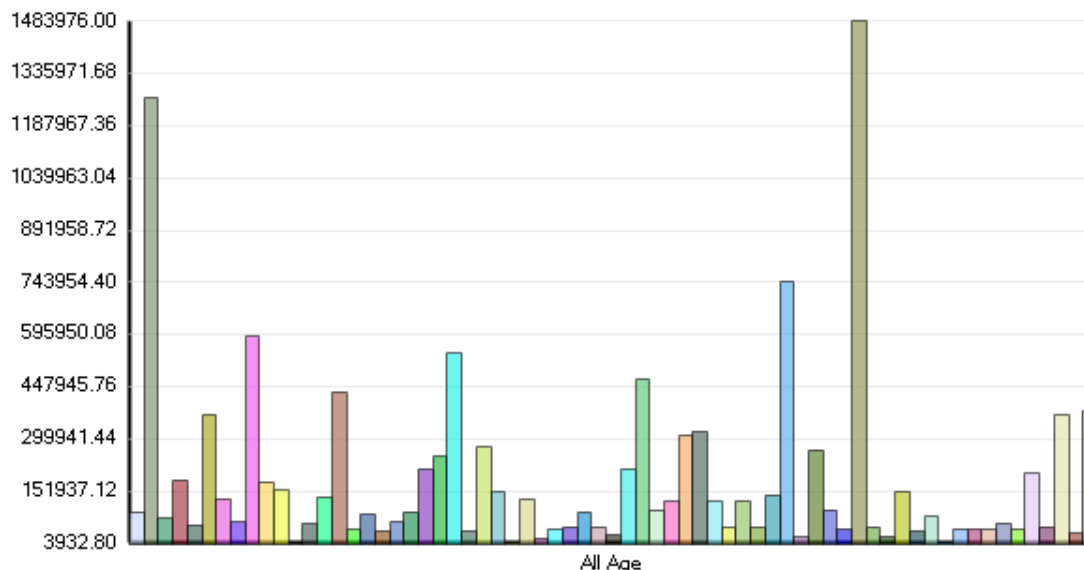


Dimensions

Birthweight Diagnosis **Educ**

- ☐ Complications of pregnanc
- ☐ Congenital anomalies
- ☐ Endocrine, nutritional and
- ☐ External causes of injury a
- ☐ Infectious and parasitic dis
- ☐ Injury and Poisoning
- ☐ Mental disorders
- ☐ Neoplasms
- ☐ Symptoms, signs, and ill-d
- ☐ the circulatory system
 - ☐ Acute rheumatic fever
 - ☐ arteries, arterioles, an
 - ☐ Cerebrovascular disea
 - ☐ Chronic rheumatic hea
 - ☐ Hypertensive disease
 - ☒ Ischemic heart diseases
 - ☐ Other forms of heart di

PA Population



- Adams_2000_Ischemic heart
- Allegheny_2000_Ischemic h
- Armstrong_2000_Ischemic r
- Beaver_2000_Ischemic hea
- Bedford_2000_Ischemic hea
- Berks_2000_Ischemic heart
- Blair_2000_Ischemic heart c
- Bradford_2000_Ischemic he
- Bucks_2000_Ischemic heart
- Butler_2000_Ischemic heart
- Cambria_2000_Ischemic he
- Cameron_2000_Ischemic he
- Carbon_2000_Ischemic hea
- Centre_2000_Ischemic hear
- Chester_2000_Ischemic hea

Row

Column

Age

Geography
Year

Background

- Birthweight
- Diagnosis
- Education
- Measures
- Race
- Region
- Sex

Query

Top 5

Compare

Reset

All Age_2000_Ischemic heart disease

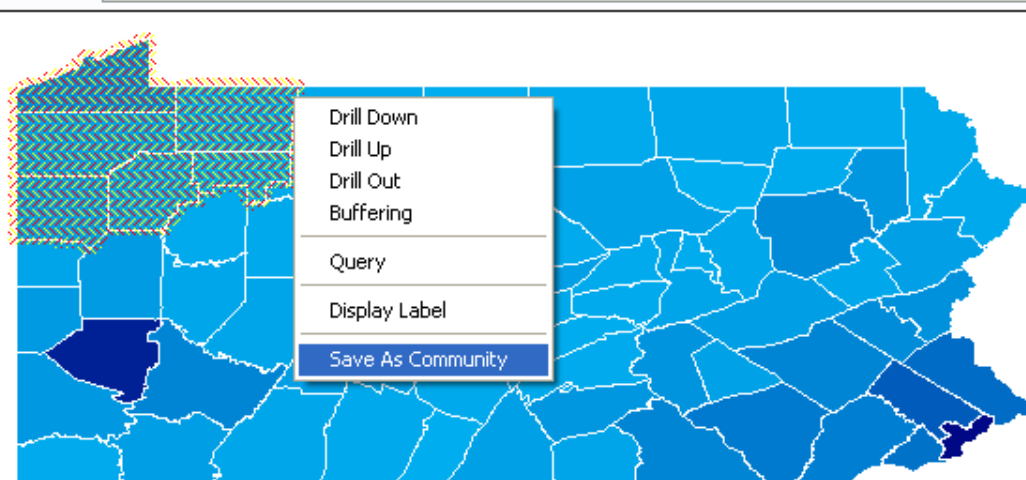
- ☐ ☒ Counties
 - 0
 - 320000
 - 630000
 - 940000
 - 1250000
 - 1560000
- ☐ ☒ buffer

Full Extend

Select Mode

Zoom Mode

Drag Mode



Dimensions

Birthweight Diagnosis Educat

- Complications of pregnanc
- Conegenital anomalies
- Endocrine, nutritional and
- External causes of injury a
- Infectious and parasitic dis
- Injury and Poisoning
- Mental disorders
- Neoplasms
- Symptoms, signs, and ill-d
- the circulatory system
 - Acute rheumatic fever
 - arteries, arterioles, an
 - Cerebrovascular disea
 - Chronic rheumatic hea
 - Hypertensive disease
 - Ischemic heart diseases
 - Other forms of heart di

Row

Column

Age

Geography
Year

Background

- Birthweight
- Diagnosis
- Education
- Measures
- Race
- Region
- Sex

Query

Top 5

Compare

Reset

PA Population

1483976.00
1335971.68
1187967.36
1039963.04
891958.72
743954.40
595950.08
447945.76
299941.44
151937.12
3932.80

Creating New Community

Enter the name of the community

OK

Cancel

NW PA

All Age_2000_Is

- ☒ Counties
 - 0
 - 320000
 - 630000
 - 940000
 - 1250000
 - 1560000

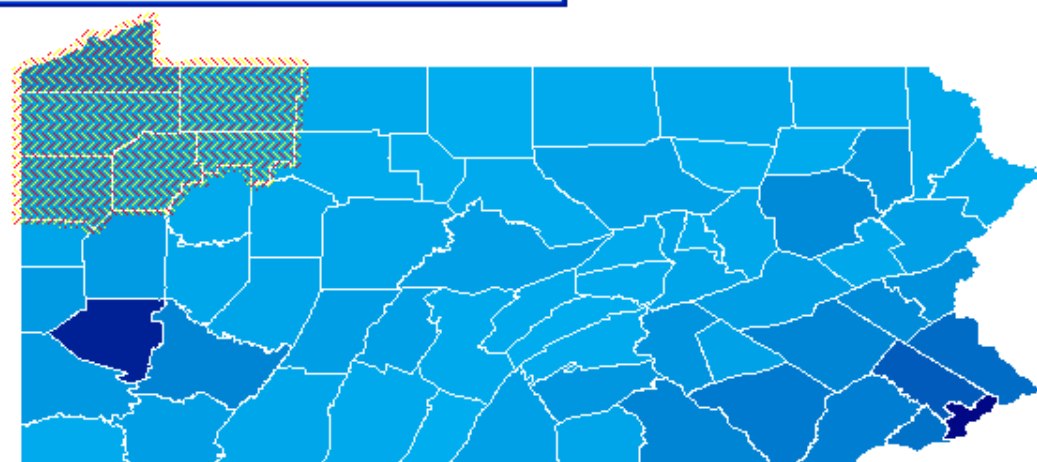
☒ buffer

Full Extend

Select Mode

Zoom Mode

Drag Mode



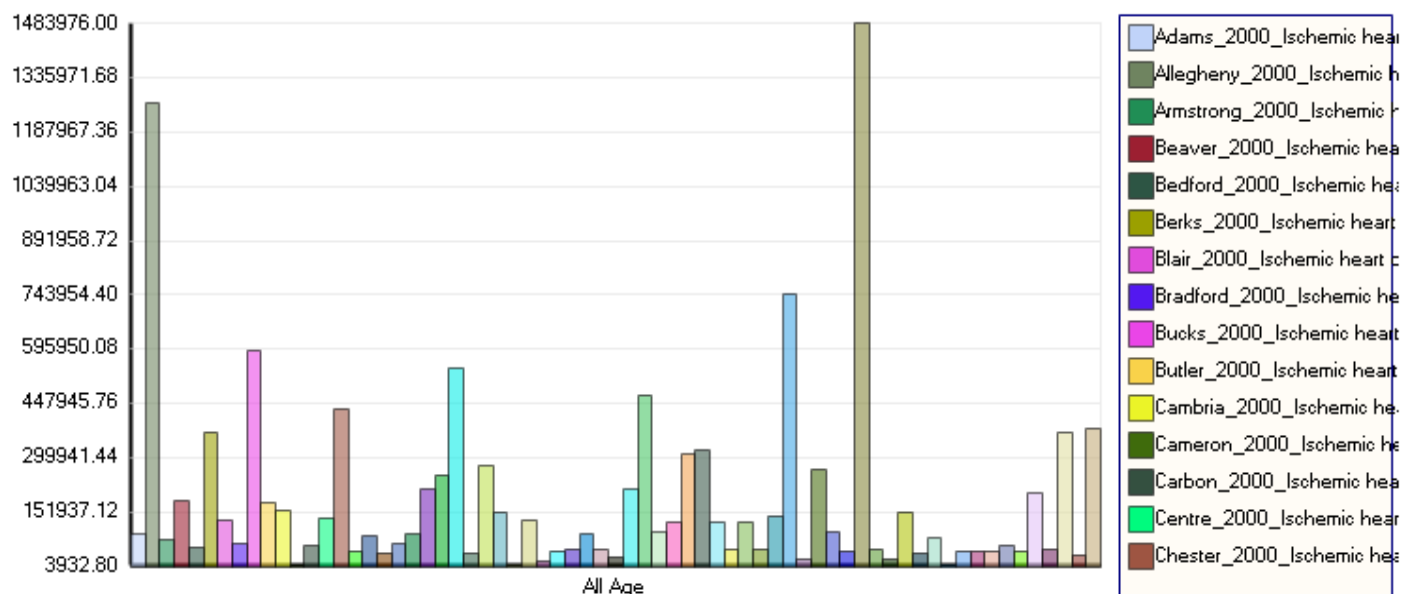
- Adams_2000_Ischemic hear
- Allegheny_2000_Ischemic h
- Armstrong_2000_Ischemic r
- Beaver_2000_Ischemic hea
- Bedford_2000_Ischemic hea
- Berks_2000_Ischemic heart
- Blair_2000_Ischemic heart c
- Bradford_2000_Ischemic he
- Bucks_2000_Ischemic heart
- Butler_2000_Ischemic heart
- Cambria_2000_Ischemic he
- Cameron_2000_Ischemic he
- Carbon_2000_Ischemic hea
- Centre_2000_Ischemic hear
- Chester_2000_Ischemic hea

Dimensions

Sex Year Community

- Counties
 - All PA Counties
 - NW PA
- Municipalities

PA Population



Row

Column

Age

Geography
Year

Background

- Birthweight
- Diagnosis
- Education
- Measures
- Race
- Region
- Sex

Query

Top 5

Compare

Reset

All Age_2000_Ischemic heart disease

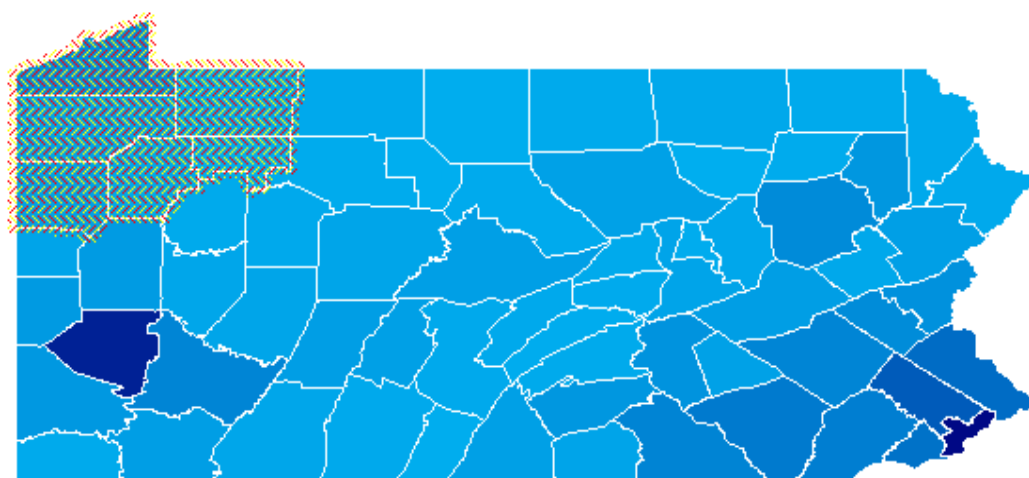
- ☒ Counties
 - 0
 - 320000
 - 630000
 - 940000
 - 1250000
 - 1560000
- ☒ buffer

Full Extend

Select Mode

Zoom Mode

Drag Mode



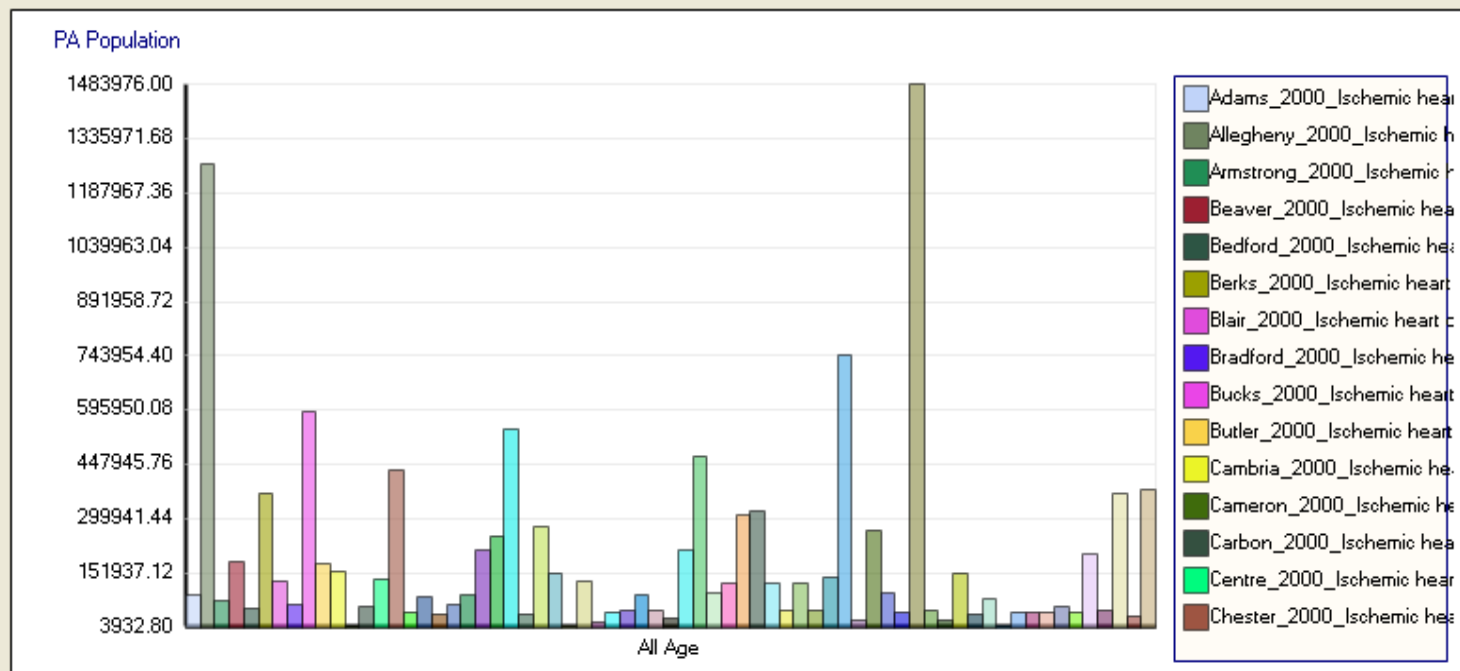
Dimensions

Sex Year Community

Counties

- All PA Counties
- NW PA
- NE PA

Municipalities



Row

Age

Column

Geography
Year

Background

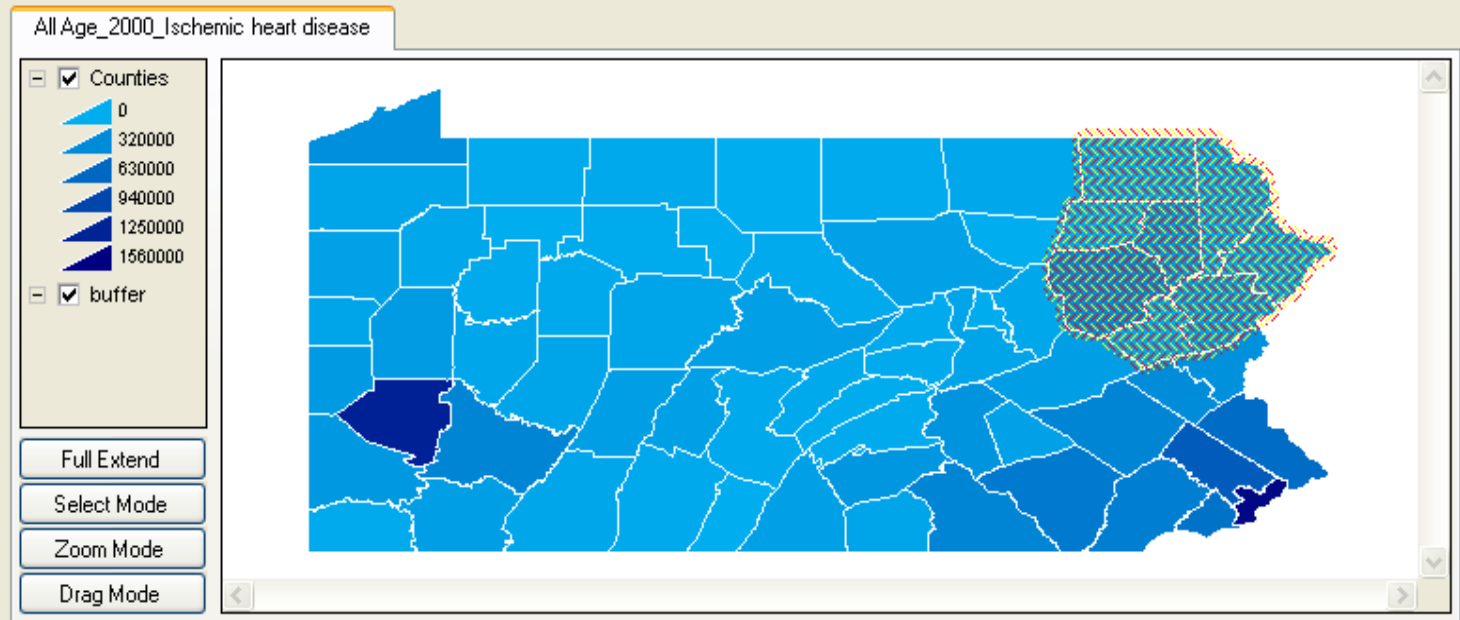
- Birthweight
- Diagnosis
- Education
- Measures
- Race
- Region
- Sex

Query

Top 5

Compare

Reset



Dimensions

Sex Year **Community**

- [-] Counties
 - All PA Counties
 - NW PA
 - NE PA
- [+] Municipalities

PA Population

886534.00
845235.56
803937.12
762638.68
721340.24
680041.80
638743.36
597444.92
556146.48
514848.04



NW PA_2000_Ischemic
NE PA_2000_Ischemic

All Age

Row

Column

Age

Geography
Year

Background

- Birthweight
- Diagnosis
- Education
- Measures
- Race
- Region
- Sex

Query

Top 5

Compare

Reset

All Age_2000_Ischemic heart disease

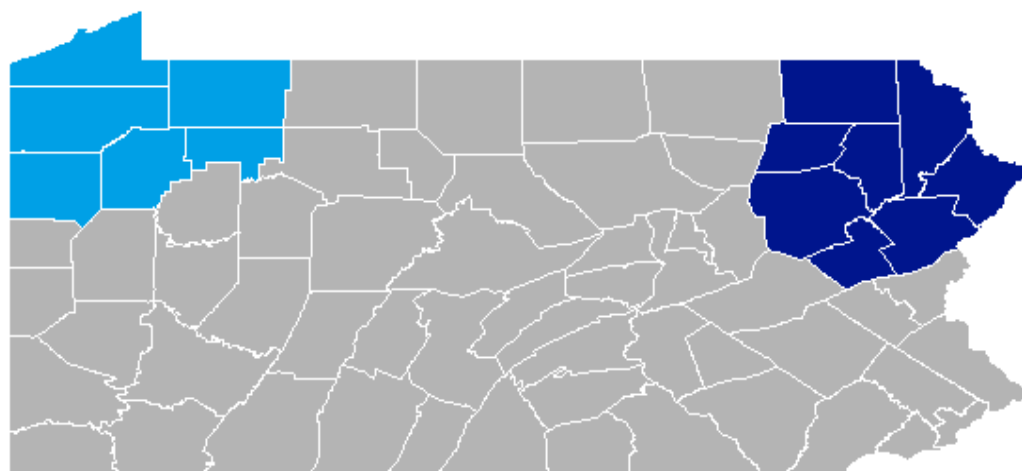
- ☒ Counties
 - 560000
 - 640000
 - 710000
 - 780000
 - 860000
 - 930000
- ☒ buffer

Full Extend

Select Mode

Zoom Mode

Drag Mode

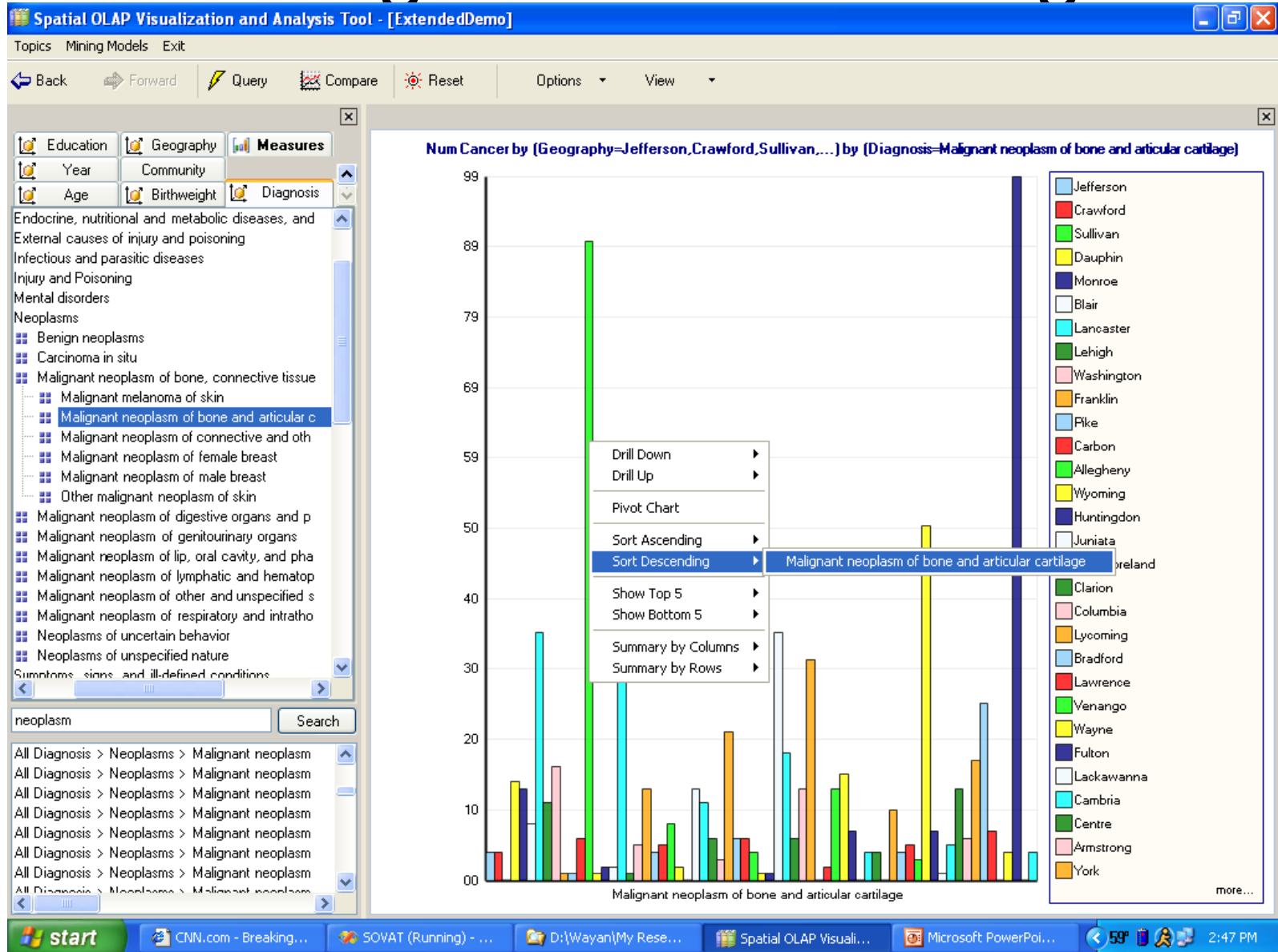




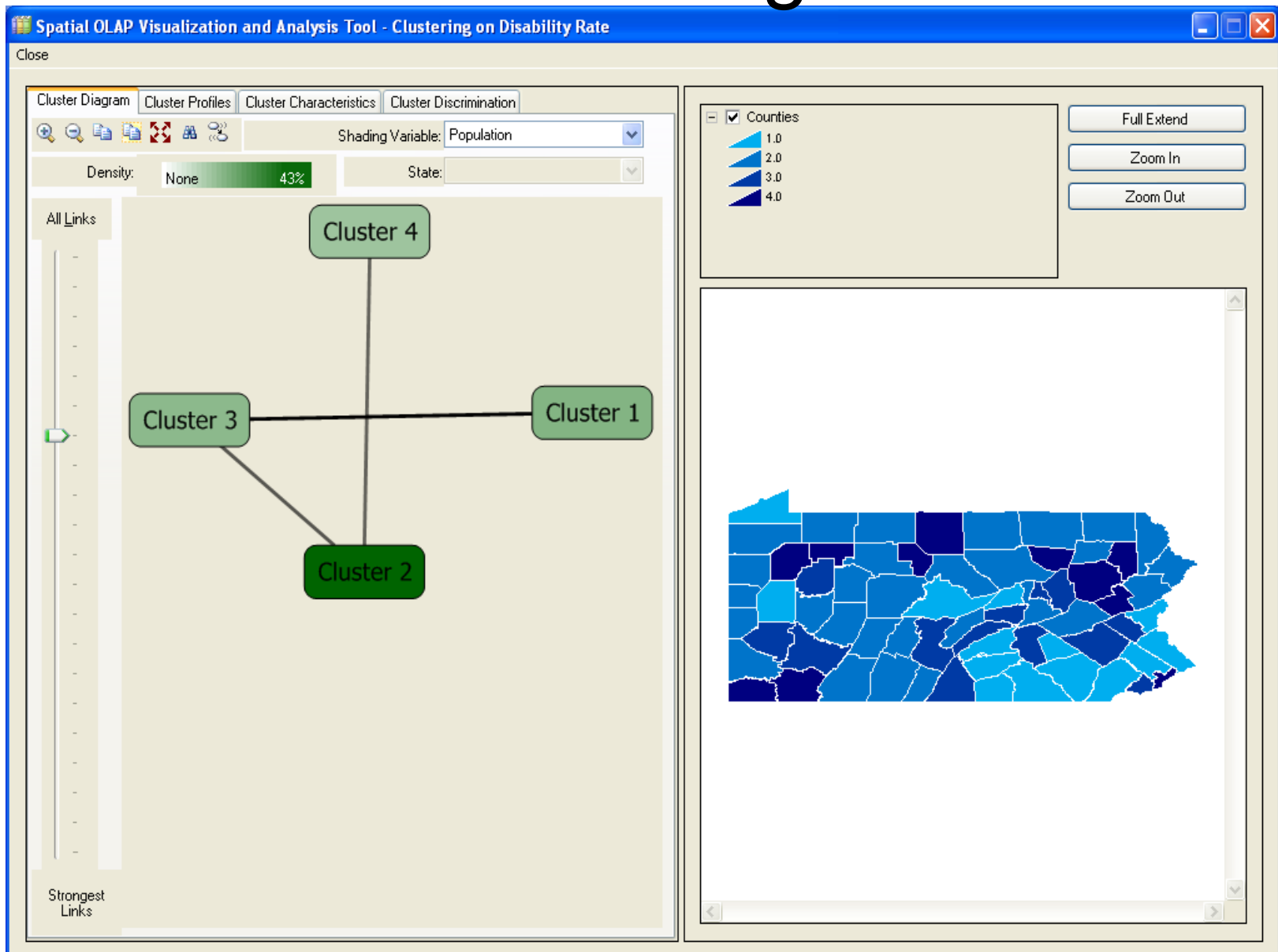
Additional Functionality



Search large dimension: diagnosis



Clustering





Example for environmental health



Number of Cancer Incident in Allegheny Counties (Municipalities level)

Topics Mining Models Exit

Back Forward Query Compare Reset Options View Buffering

Diagnosis Education Geography

Age Allegheny... Birthweight

Measures Race Region

Birth/100,000 (Smoothed)

Cancer Incidence/100,000

Cancer Incident/100,000 (Smoothed)

Death/100,000

Death/100,000 (Smoothed)

Gamma Cancer

Gamma Death

Inpatient Charges

Inpatient/1,000

LDS

LDS/1,000

Num Births

Num Cancer

Num Deaths

Num Inpatient

Num Outpatient

Outpatient Charges

Outpatient/1,000

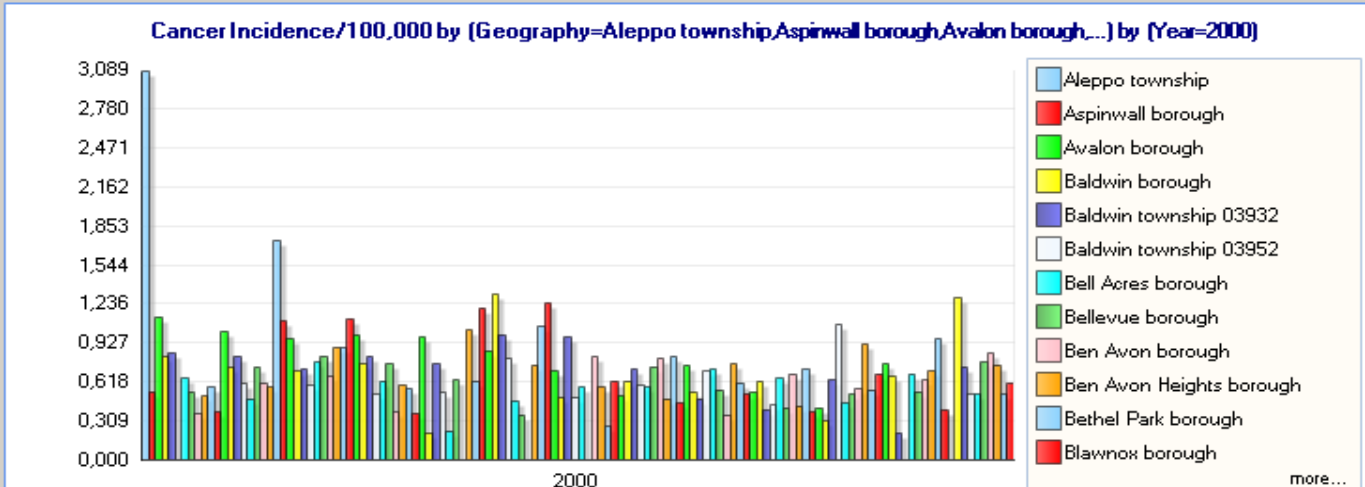
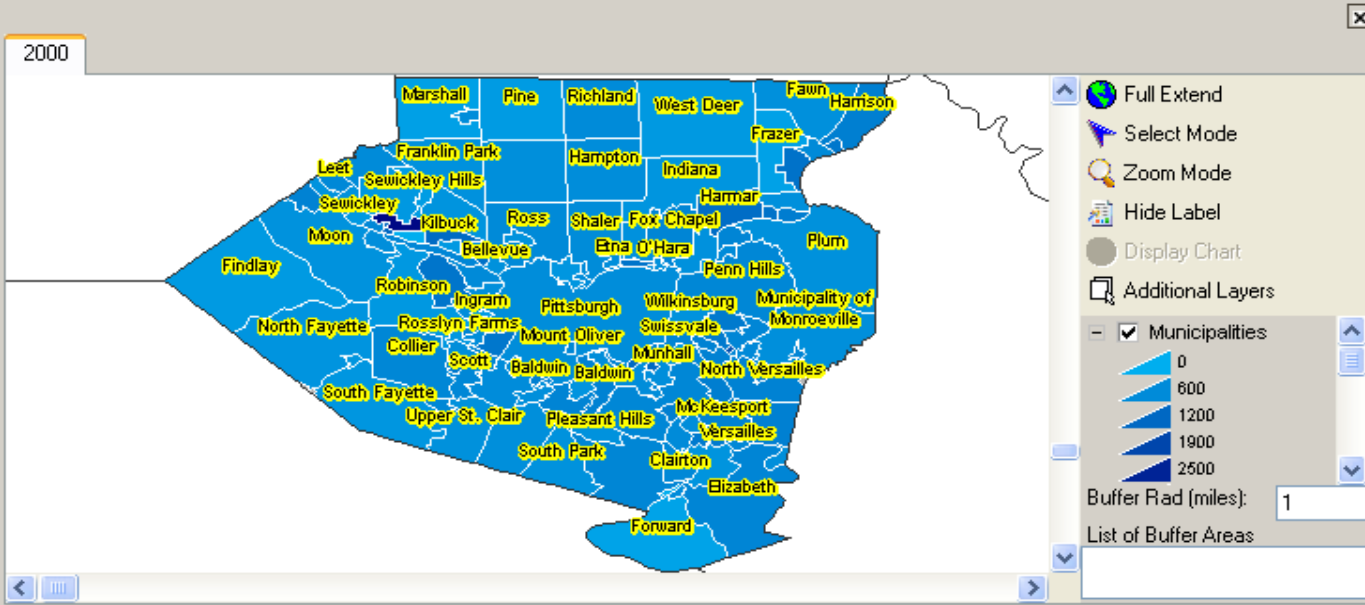
Population

Population - AlleghenyPopulation

Variance Birth

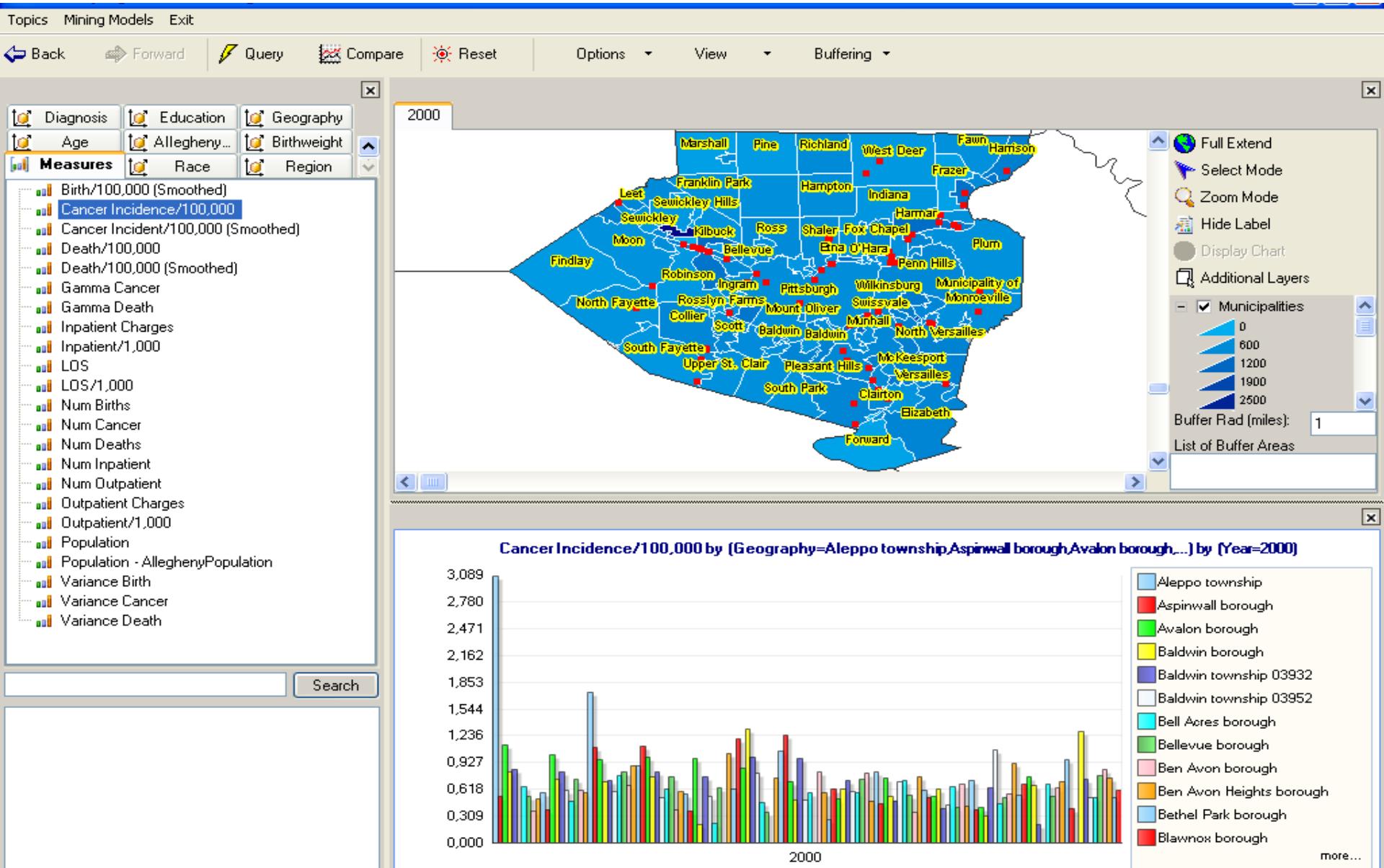
Variance Cancer

Variance Death



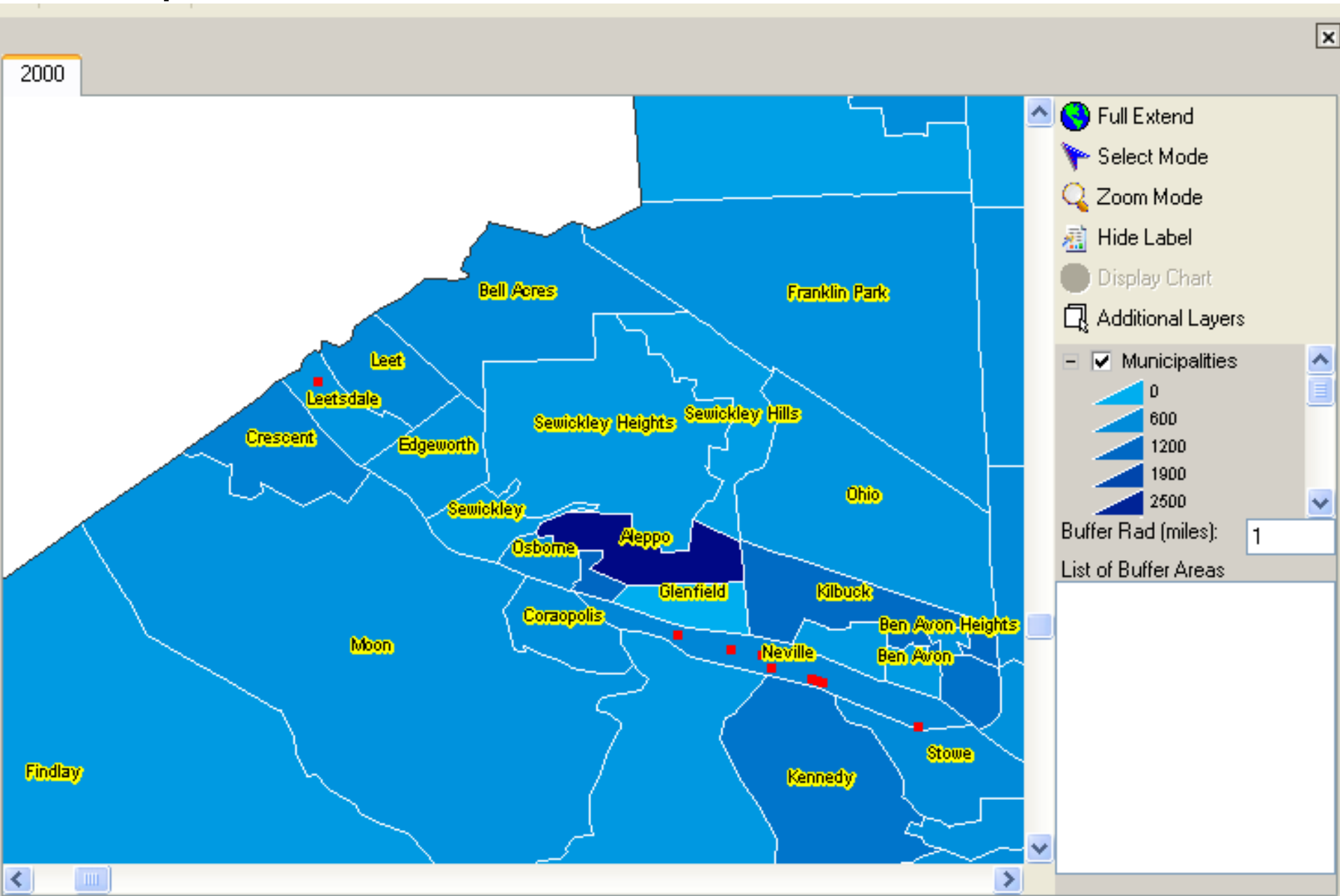


Red nodes shows industrial places in Allegheny County



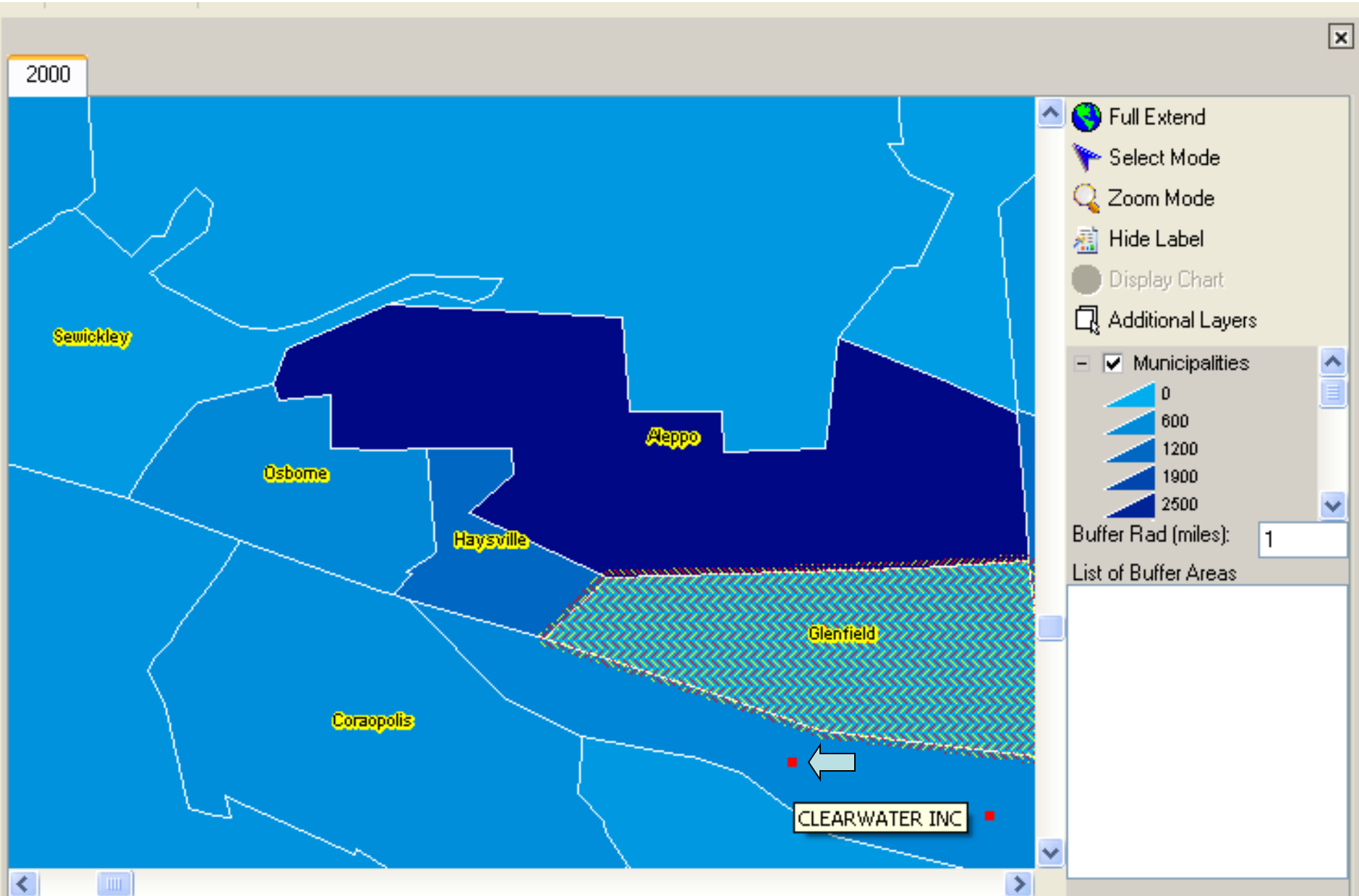


A close up view of the nodes



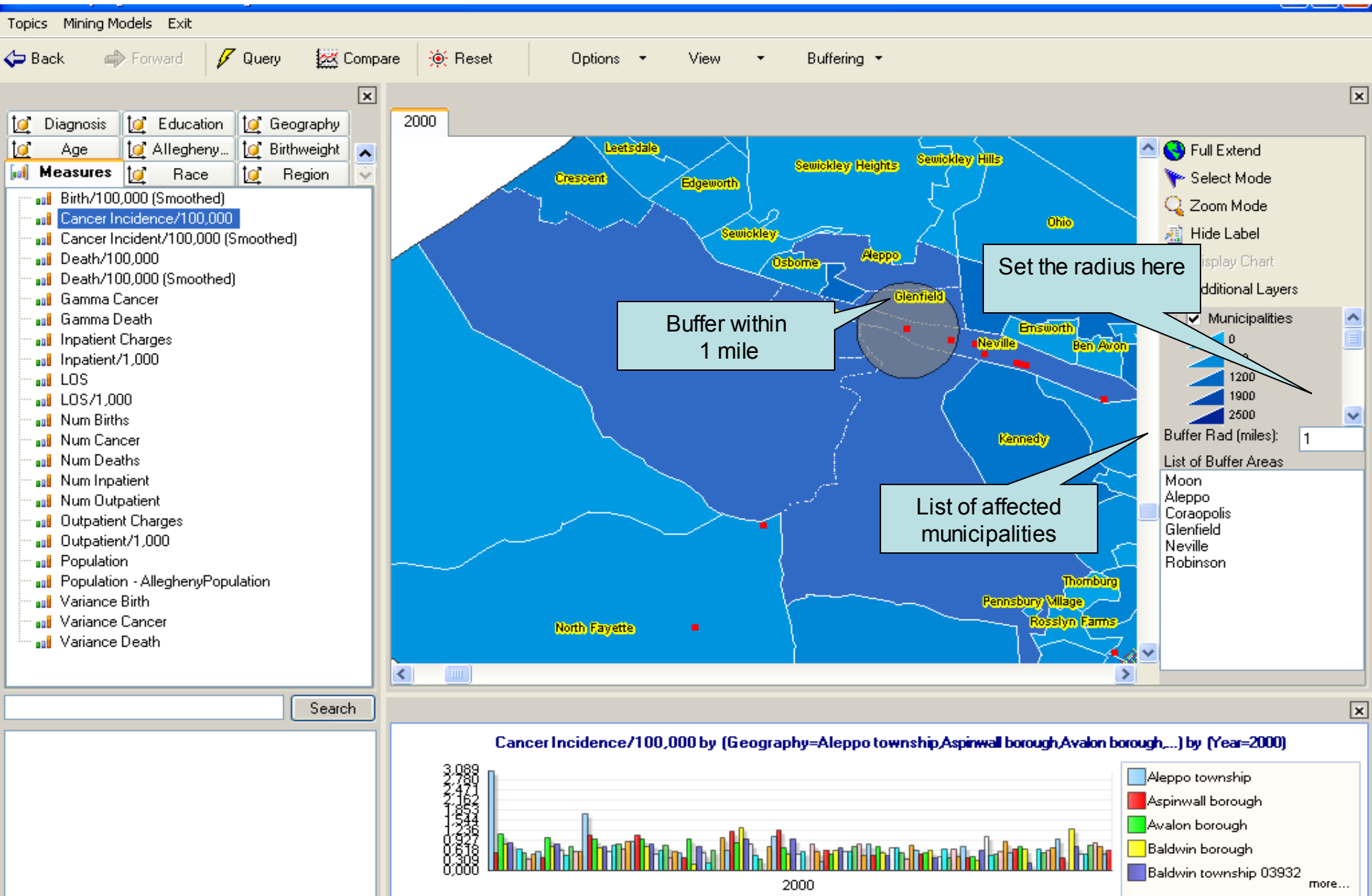


The arrow shows an industrial point : CLEARWATER INC





Buffer within 1 mile from CLEARWATER INC and the affected municipalities





Buffer within 2.5 mile from CLEARWATER INC and the affected municipalities

Topics Mining Models Exit

Back Forward Query Compare Reset Options View Buffering

Diagnosis Education Geography

Age Allegheny... Birthweight

Measures Race Region

Birth/100,000 (Smoothed)

Cancer Incidence/100,000

Cancer Incident/100,000 (Smoothed)

Death/100,000

Death/100,000 (Smoothed)

Gamma Cancer

Gamma Death

Inpatient Charges

Inpatient/1,000

LOS

LOS/1,000

Num Births

Num Cancer

Num Deaths

Num Inpatient

Num Outpatient

Outpatient Charges

Outpatient/1,000

Population

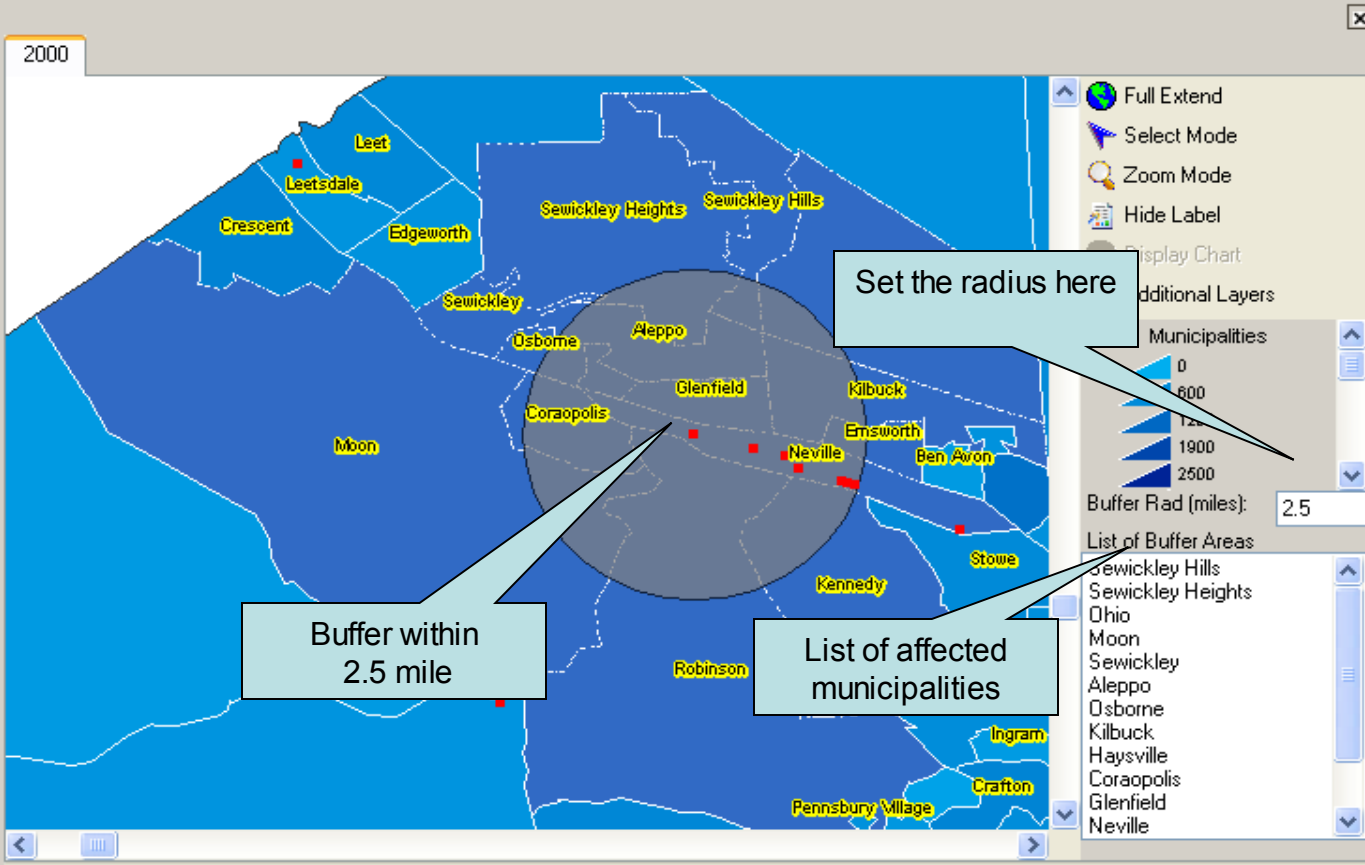
Population - AlleghenyPopulation

Variance Birth

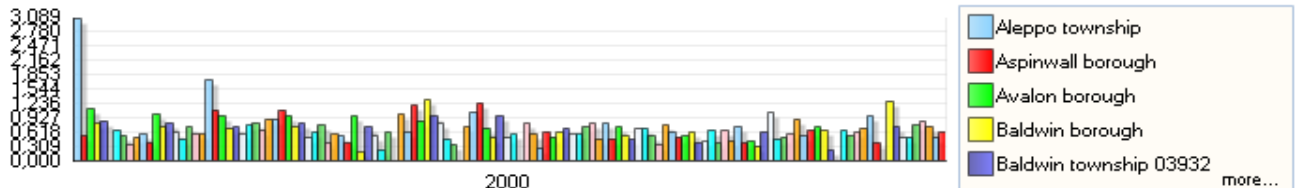
Variance Cancer

Variance Death

Search



Cancer Incidence/100,000 by (Geography=Aleppo township,Aspinwall borough,Avalon borough,...) by (Year=2000)





Using Other Data Sets

